



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>(51) International Patent Classification <sup>6</sup> :<br/>C12N 15/31, C07K 14/35, A61K 39/04,<br/>48/00, 49/00, C12N 15/62, C07K 19/00,<br/>G01N 33/50, 33/60, 33/569, C12N 1/19,<br/>1/20, 1/21, 5/10 // (C12N 1/21, C12N<br/>1:19)</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | A2 | <p>(11) International Publication Number: <b>WO 98/53075</b></p> <p>(43) International Publication Date: 26 November 1998 (26.11.98)</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <p>(21) International Application Number: PCT/US98/10407</p> <p>(22) International Filing Date: 20 May 1998 (20.05.98)</p> <p>(30) Priority Data:<br/>08/859,381 20 May 1997 (20.05.97) US<br/>09/073,016 5 May 1998 (05.05.98) US</p> <p>(71) Applicant: CORIXA CORPORATION [US/US]; Suite 200,<br/>1124 Columbia Street, Seattle, WA 98104 (US).</p> <p>(72) Inventors: ALDERSON, Mark, R.; 1116 Grove Avenue<br/>Northwest, Bainbridge Island, WA 98110 (US); DILLON,<br/>David, C.; 21607 Northeast 24th Street, Redmond, WA<br/>98053 (US); SKIEKY, Yassir, A. W.; #327 - 25th<br/>Avenue N.W., Seattle, WA 98117 (US); CAMPOS-NETO,<br/>Antonio; 9308 N.E. Midship Court, Bainbridge Island, WA<br/>98110 (US).</p> <p>(74) Agents: MAKI, David, J. et al.; Seed and Berry LLP,<br/>6300 Columbia Center, 701 Fifth Avenue, Seattle, WA<br/>98104-7092 (US).</p> |    | <p>(81) Designated States: AL, AM, AT, AU, BA, BB, BG, BR, BY,<br/>CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GR,<br/>GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC,<br/>LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX,<br/>NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TH,<br/>TM, TR, TT, UA, UG, UZ, VN, YD, ZW, ARIPO patent<br/>(GM, GM, KE, LS, MW, SD, SZ, DZ, ZM), European patent<br/>(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent<br/>(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT,<br/>LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI,<br/>CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p><b>Published</b><br/><i>Without international search report and to be republished<br/>upon receipt of that report.</i></p> |
| <p>(84) Title: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF TUBERCULOSIS AND METHODS OF THEIR USE</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <p>(87) Abstract</p> <p>Compounds and methods for inducing protective immunity against tuberculosis are disclosed. The compounds provided include polypeptides that contain at least one immunogenic portion of one or more <i>M. tuberculosis</i> proteins and DNA molecules encoding such polypeptides. Such compounds may be formulated into vaccines and/or pharmaceutical compositions for immunization against <i>M. tuberculosis</i> infection, or may be used for the diagnosis of tuberculosis.</p>                                                                                                                                                                                                                                                                                                                                                           |    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

# FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |                                       |    |                                           |    |                          |
|----|--------------------------|----|---------------------------------------|----|-------------------------------------------|----|--------------------------|
| AL | Albania                  | BN | Brunei                                | LS | Lesotho                                   | SI | Slovenia                 |
| AM | Armenia                  | FI | Finland                               | LT | Lithuania                                 | SK | Slovakia                 |
| AT | Austria                  | FR | France                                | LU | Luxembourg                                | SN | Senegal                  |
| AU | Australia                | GA | Gabon                                 | LV | Latvia                                    | SE | Sweden                   |
| AX | Azerbaijan               | GB | United Kingdom                        | MC | Monaco                                    | TD | Chad                     |
| BA | Bosnia and Herzegovina   | GE | Georgia                               | MD | Republic of Moldova                       | TG | Togo                     |
| BB | Barbados                 | GR | Greece                                | MG | Madagascar                                | TJ | Tajikistan               |
| BE | Belgium                  | GN | Guinea                                | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan             |
| BF | Burkina Faso             | HR | Croatia                               | ML | Mali                                      | TR | Turkey                   |
| BG | Bulgaria                 | HU | Hungary                               | MM | Myanmar                                   | TT | Trinidad and Tobago      |
| BJ | Benin                    | IE | Ireland                               | MN | Mongolia                                  | UA | Ukraine                  |
| BR | Brazil                   | IL | Israel                                | MR | Mauritania                                | UG | Uganda                   |
| BV | Belarus                  | IS | Iceland                               | MT | Malta                                     | US | United States of America |
| CA | Canada                   | IT | Italy                                 | MX | Mexico                                    | UZ | Uzbekistan               |
| CF | Central African Republic | JP | Japan                                 | NE | Niger                                     | VN | Viet Nam                 |
| CG | Congo                    | KE | Kenya                                 | NL | Netherlands                               | YE | Yemen                    |
| CH | Switzerland              | KG | Kyrgyzstan                            | NO | Norway                                    | ZW | Zimbabwe                 |
| CI | Cote d'Ivoire            | KP | Democratic People's Republic of Korea | NZ | New Zealand                               |    |                          |
| CM | Cameroon                 | KR | Republic of Korea                     | PL | Poland                                    |    |                          |
| CN | China                    | KW | Kuwait                                | PT | Portugal                                  |    |                          |
| CZ | Czech Republic           | LA | Laos                                  | RO | Romania                                   |    |                          |
| DE | Germany                  | LC | Saint Lucia                           | RU | Russian Federation                        |    |                          |
| DK | Denmark                  | LI | Liechtenstein                         | SD | Sudan                                     |    |                          |
| EE | Estonia                  | LK | Sri Lanka                             | SE | Sweden                                    |    |                          |
|    |                          | LR | Liberia                               | SG | Singapore                                 |    |                          |

DescriptionCOMPOUNDS FOR IMMUNOTHERAPY AND  
DIAGNOSIS OF TUBERCULOSIS AND METHODS OF THEIR USE5    Technical Field

The present invention relates generally to detecting, treating and preventing *Mycobacterium tuberculosis* infection. The invention is more particularly related to polypeptides comprising a *Mycobacterium tuberculosis* antigen, or a portion or other variant thereof, and the use of such polypeptides for diagnosing and vaccinating  
10    against *Mycobacterium tuberculosis* infection.

Background of the Invention

Tuberculosis is a chronic, infectious disease, that is generally caused by infection with *Mycobacterium tuberculosis*. It is a major disease in developing  
15    countries, as well as an increasing problem in developed areas of the world, with about 8 million new cases and 3 million deaths each year. Although the infection may be asymptomatic for a considerable period of time, the disease is most commonly manifested as an acute inflammation of the lungs, resulting in fever and a nonproductive cough. If left untreated, serious complications and death typically result.

20    Although tuberculosis can generally be controlled using extended antibiotic therapy, such treatment is not sufficient to prevent the spread of the disease. Infected individuals may be asymptomatic, but contagious, for some time. In addition, although compliance with the treatment regimen is critical, patient behavior is difficult to monitor. Some patients do not complete the course of treatment, which can lead to  
25    ineffective treatment and the development of drug resistance.

Inhibiting the spread of tuberculosis requires effective vaccination and accurate, early diagnosis of the disease. Currently, vaccination with live bacteria is the most efficient method for inducing protective immunity. The most common *Mycobacterium* employed for this purpose is *Bacillus Calmette-Guerin* (BCG), an  
30    avirulent strain of *Mycobacterium bovis*. However, the safety and efficacy of BCG is a source of controversy and some countries, such as the United States, do not vaccinate

the general public. Diagnosis is commonly achieved using a skin test, which involves intradermal exposure to tuberculin PPD (protein-purified derivative). Antigen-specific T cell responses result in measurable induration at the injection site by 48-72 hours after injection, which indicates exposure to Mycobacterial antigens. Sensitivity and specificity have, however, been a problem with this test, and individuals vaccinated with BCG cannot be distinguished from infected individuals.

While macrophages have been shown to act as the principal effectors of *M. tuberculosis* immunity, T cells are the predominant inducers of such immunity. The essential role of T cells in protection against *M. tuberculosis* infection is illustrated by the frequent occurrence of *M. tuberculosis* in AIDS patients, due to the depletion of CD4 T cells associated with human immunodeficiency virus (HIV) infection. Mycobacterium-reactive CD4 T cells have been shown to be potent producers of gamma-interferon (IFN- $\gamma$ ), which, in turn, has been shown to trigger the antimycobacterial effects of macrophages in mice. While the role of IFN- $\gamma$  in humans is less clear, studies have shown that 1,25-dihydroxy-vitamin D3, either alone or in combination with IFN- $\gamma$  or tumor necrosis factor-alpha, activates human macrophages to inhibit *M. tuberculosis* infection. Furthermore, it is known that IFN- $\gamma$  stimulates human macrophages to make 1,25-dihydroxy-vitamin D3. Similarly, IL-12 has been shown to play a role in stimulating resistance to *M. tuberculosis* infection. For a review of the immunology of *M. tuberculosis* infection see Chan and Kaufmann in *Tuberculosis. Pathogenesis, Protection and Control*, Bloom (ed.), ASM Press, Washington, DC, 1994.

Accordingly, there is a need in the art for improved vaccines and methods for preventing, treating and detecting tuberculosis. The present invention fulfills these needs and further provides other related advantages.

#### Summary of the Invention

Briefly stated, this invention provides compounds and methods for preventing and diagnosing tuberculosis. In one aspect, polypeptides are provided comprising an immunogenic portion of an *M. tuberculosis* antigen, or a variant of such an antigen that differs only in conservative substitutions and/or modifications, the

antigen comprising an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID NO: 1, 11, 12, 83, 103-108, 125, 127, 129-137, 139 and 140, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NO: 1, 11, 12, 83, 103-108, 125, 127, 129-137, 139 and 140, or a complement thereof under moderately stringent conditions. In a second aspect, the present invention provides polypeptides comprising an immunogenic portion of a *M. tuberculosis* antigen having an amino acid sequence selected from the group consisting of sequences provided in SEQ ID NO: 16-33, 109, 126, 138, 141, 142 and variants thereof.

10 In related aspects, DNA sequences encoding the above polypeptides, expression vectors comprising these DNA sequences and host cells transformed or transfected with such expression vectors are also provided.

In another aspect, the present invention provides fusion proteins comprising a first and a second inventive polypeptide or, alternatively, an inventive polypeptide and a known *M. tuberculosis* antigen.

15 Within other aspects, the present invention provides pharmaceutical compositions that comprise one or more of the above polypeptides, or a DNA molecule encoding such polypeptides, and a physiologically acceptable carrier. The invention also provides vaccines comprising one or more of the polypeptides as described above and a non-specific immune response enhancer, together with vaccines comprising one or more DNA sequences encoding such polypeptides and a non-specific immune response enhancer.

20 In yet another aspect, methods are provided for inducing protective immunity in a patient, comprising administering to a patient an effective amount of one or more of the above polypeptides.

In further aspects of this invention, methods and diagnostic kits are provided for detecting tuberculosis in a patient. The methods comprise contacting dermal cells of a patient with one or more of the above polypeptides and detecting an immune response on the patient's skin. The diagnostic kits comprise one or more of the

above polypeptides in combination with an apparatus sufficient to contact the polypeptide with the dermal cells of a patient.

In yet another aspect, methods are provided for detecting tuberculosis in a patient, such methods comprising contacting dermal cells of a patient with one or  
5 more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID NO: 2-10, 102, 128, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NO: 2-10, 102, 128; and detecting an immune response on the patient's skin. Diagnostic kits for use in such methods are also provided.

10 These and other aspects of the present invention will become apparent upon reference to the following detailed description and attached drawings. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

#### 15 Brief Description of the Drawings

Figures 1A and 1B illustrate the stimulation of proliferation and interferon- $\gamma$  production, respectively, in T cells derived from a first PPD-positive donor (referred to as D7) by recombinant ORF-2 and synthetic peptides to ORF-2.

20 Figures 2A and 2B illustrate the stimulation of proliferation and interferon- $\gamma$  production, respectively, in T cells derived from a second PPD-positive donor (referred to as D160) by recombinant ORF-2 and synthetic peptides to ORF-2.

#### Detailed Description of the Invention

As noted above, the present invention is generally directed to  
25 compositions and methods for preventing, treating and diagnosing tuberculosis. The compositions of the subject invention include polypeptides that comprise at least one immunogenic portion of a *M. tuberculosis* antigen, or a variant of such an antigen that differs only in conservative substitutions and/or modifications. As used herein, the term "polypeptide" encompasses amino acid chains of any length, including full length  
30 proteins (i.e., antigens), wherein the amino acid residues are linked by covalent peptide bonds. Thus, a polypeptide comprising an immunogenic portion of one of the above

antigens may consist entirely of the immunogenic portion, or may contain additional sequences. The additional sequences may be derived from the native *M. tuberculosis* antigen or may be heterologous, and such sequences may (but need not) be immunogenic.

9 "Immunogenic," as used herein, refers to the ability to elicit an immune response (e.g., cellular) in a patient, such as a human, and/or in a biological sample. In particular, antigens that are immunogenic (and immunogenic portions or other variants of such antigens) are capable of stimulating cell proliferation, interleukin-12 production and/or interferon- $\gamma$  production in biological samples comprising one or more cells  
10 selected from the group of T cells, NK cells, B cells and macrophages, where the cells are derived from an *M. tuberculosis*-immune individual. Polypeptides comprising at least an immunogenic portion of one or more *M. tuberculosis* antigens may generally be used to detect tuberculosis or to induce protective immunity against tuberculosis in a patient.

15 The compositions and methods of this invention also encompass variants of the above polypeptides. A polypeptide "variant," as used herein, is a polypeptide that differs from the recited polypeptide only in conservative substitutions and/or modifications, such that the therapeutic, antigenic and/or immunogenic properties of the polypeptide are retained. Polypeptide variants preferably exhibit at least about 70%,  
20 more preferably at least about 90% and most preferably at least about 95% identity to the identified polypeptides. For polypeptides with immunoreactive properties, variants may, alternatively, be identified by modifying the amino acid sequence of one of the above polypeptides, and evaluating the immunoreactivity of the modified polypeptide. For polypeptides useful for the generation of diagnostic binding agents, a variant may  
25 be identified by evaluating a modified polypeptide for the ability to generate antibodies that detect the presence or absence of tuberculosis. Alternatively, variants of the claimed antigens that may be usefully employed in the inventive diagnostic methods may be identified by evaluating modified polypeptides for their ability to detect antibodies present in the sera of tuberculosis-infected patients. Such modified

sequences may be prepared and tested using, for example, the representative procedures described herein.

A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydrophobic nature of the polypeptide to be substantially unchanged. In general, the following groups of amino acids represent conservative changes: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, scr, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his, and (5) phe, tyr, trp, his.

Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenic properties, secondary structure and hydrophobic nature of the polypeptide. For example, a polypeptide may be conjugated to a signal (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

20

In general, *M. tuberculosis* antigens, and DNA sequences encoding such antigens, may be prepared using any of a variety of procedures. For example, genomic or cDNA libraries derived from *M. tuberculosis* may be screened directly using peripheral blood mononuclear cells (PBMCs) or T cell lines or clones derived from one or more *M. tuberculosis*-immune individuals. Direct library screens may generally be performed by assaying pools of expressed recombinant proteins for the ability of induce proliferation and/or interferon- $\gamma$  production in T cells derived from an *M. tuberculosis*-immune individual. Potential T cell antigens may be first selected based on antibody reactivity, as described above.

Alternatively, DNA sequences encoding antigens may be identified by screening an appropriate *M. tuberculosis* genomic or cDNA expression library with sera obtained from patients infected with *M. tuberculosis*. Such screens may generally be performed using techniques well known to those of ordinary skill in the art, such as those described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989.

Purified antigens are then evaluated for their ability to elicit an appropriate immune response (e.g., cellular) using, for example, the representative methods described herein. Immunogenic antigens may then be partially sequenced using techniques such as traditional Edman chemistry. See Edman and Berg, *Eur. J. Biochem.* 80:116-132, 1967. Immunogenic antigens may also be produced recombinantly using a DNA sequence that encodes the antigen, which has been inserted into an expression vector and expressed in an appropriate host.

DNA sequences encoding the inventive antigens may also be obtained by screening an appropriate *M. tuberculosis* cDNA or genomic DNA library for DNA sequences that hybridize to degenerate oligonucleotides derived from partial amino acid sequences of isolated antigens. Degenerate oligonucleotide sequences for use in such a screen may be designed and synthesized, and the screen may be performed, as described (for example) in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989 (and references cited therein). Polymerase chain reaction (PCR) may also be employed, using the above oligonucleotides in methods well known in the art, to isolate a nucleic acid probe from a cDNA or genomic library. The library screen may then be performed using the isolated probe.

Regardless of the method of preparation, the antigens (and immunogenic portions thereof) described herein have the ability to induce an immunogenic response. More specifically, the antigens have the ability to induce proliferation and/or cytokine production (i.e., interferon- $\gamma$  and/or interleukin-12 production) in T cells, NK cells, B cells and/or macrophages derived from an *M. tuberculosis*-immune individual. The selection of cell type for use in evaluating an immunogenic response to a antigen will,

of course, depend on the desired response. For example, interleukin-12 production is most readily evaluated using preparations containing B cells and/or macrophages. An *M. tuberculosis*-immune individual is one who is considered to be resistant to the development of tuberculosis by virtue of having mounted an effective T cell response to *M. tuberculosis* (i.e., substantially free of disease symptoms). Such individuals may be identified based on a strongly positive (i.e., greater than about 10 mm diameter induration) intradermal skin test response to tuberculosis proteins (PPD) and an absence of any signs or symptoms of tuberculosis disease. T cells, NK cells, B cells and macrophages derived from *M. tuberculosis*-immune individuals may be prepared using methods known to those of ordinary skill in the art. For example, a preparation of PBMCs (i.e., peripheral blood mononuclear cells) may be employed without further separation of component cells. PBMCs may generally be prepared, for example, using density centrifugation through Ficoll™ (Winthrop Laboratories, NY).

T cells for use in the assays described herein may also be purified directly from PBMCs. Alternatively, an enriched T cell line reactive against mycobacterial proteins, or T cell clones reactive to individual mycobacterial proteins, may be employed. Such T cell clones may be generated by, for example, culturing PBMCs from *M. tuberculosis*-immune individuals with mycobacterial proteins for a period of 2-4 weeks. This allows expansion of only the mycobacterial protein-specific T cells, resulting in a line composed solely of such cells. These cells may then be cloned and tested with individual proteins, using methods known to those of ordinary skill in the art, to more accurately define individual T cell specificity. In general, antigens that test positive in assays for proliferation and/or cytokine production (i.e., interferon- $\gamma$  and/or interleukin-12 production) performed using T cells, NK cells, B cells and/or macrophages derived from an *M. tuberculosis*-immune individual are considered immunogenic. Such assays may be performed, for example, using the representative procedures described below. Immunogenic portions of such antigens may be identified using similar assays, and may be present within the polypeptides described herein.

The ability of a polypeptide (e.g., an immunogenic antigen, or a portion or other variant thereof) to induce cell proliferation is evaluated by contacting the cells

(e.g., T cells and/or NK cells) with the polypeptide and measuring the proliferation of the cells. In general, the amount of polypeptide that is sufficient for evaluation of about  $10^5$  cells ranges from about 10 ng/mL to about 100  $\mu$ g/mL and preferably is about 10  $\mu$ g/mL. The incubation of polypeptide with cells is typically performed at 37°C for about  
5 six days. Following incubation with polypeptide, the cells are assayed for a proliferative response, which may be evaluated by methods known to those of ordinary skill in the art, such as exposing cells to a pulse of radiolabeled thymidine and measuring the incorporation of label into cellular DNA. In general, a polypeptide that results in at least a three fold increase in proliferation above background (i.e., the  
10 proliferation observed for cells cultured without polypeptide) is considered to be able to induce proliferation.

The ability of a polypeptide to stimulate the production of interferon- $\gamma$  and/or interleukin-12 in cells may be evaluated by contacting the cells with the polypeptide and measuring the level of interferon- $\gamma$  or interleukin-12 produced by the  
15 cells. In general, the amount of polypeptide that is sufficient for the evaluation of about  $10^5$  cells ranges from about 10 ng/mL to about 100  $\mu$ g/mL and preferably is about 10  $\mu$ g/mL. The polypeptide may, but need not, be immobilized on a solid support, such as a bead or a biodegradable microsphere, such as those described in U.S. Patent Nos. 4,897,268 and 5,075,109. The incubation of polypeptide with the cells is typically  
20 performed at 37°C for about six days. Following incubation with polypeptide, the cells are assayed for interferon- $\gamma$  and/or interleukin-12 (or one or more subunits thereof), which may be evaluated by methods known to those of ordinary skill in the art, such as an enzyme-linked immunosorbent assay (ELISA) or, in the case of IL-12 P70 heterodimer, a bioassay such as an assay measuring proliferation of T cells. In general,  
25 a polypeptide that results in the production of at least 50 pg of interferon- $\gamma$  per mL of cultured supernatant (containing  $10^5$ - $10^6$  T cells per mL) is considered able to stimulate the production of interferon- $\gamma$ . A polypeptide that stimulates the production of at least 10 pg/mL of IL-12 P70 subunit, and/or at least 100 pg/mL of IL-12 P40 subunit, per  $10^7$  macrophages or B cells (or per  $3 \times 10^6$  PBMC) is considered able to stimulate the  
30 production of IL-12.

In general, immunogenic antigens are those antigens that stimulate proliferation and/or cytokine production (i.e., interferon- $\gamma$  and/or interleukin-12 production) in T cells, NK cells, B cells and/or macrophages derived from at least about 25% of *M. tuberculosis*-immune individuals. Among these immunogenic antigens, polypeptides having superior therapeutic properties may be distinguished based on the magnitude of the responses in the above assays and based on the percentage of individuals for which a response is observed. In addition, antigens having superior therapeutic properties will not stimulate proliferation and/or cytokine production *in vitro* in cells derived from more than about 25% of individuals that are not *M. tuberculosis*-immune, thereby eliminating responses that are not specifically due to *M. tuberculosis*-responsive cells. Those antigens that induce a response in a high percentage of T cell, NK cell, B cell and/or macrophage preparations from *M. tuberculosis*-immune individuals (with a low incidence of responses in cell preparations from other individuals) have superior therapeutic properties.

Antigens with superior therapeutic properties may also be identified based on their ability to diminish the severity of *M. tuberculosis* infection in experimental animals, when administered as a vaccine. Suitable vaccine preparations for use on experimental animals are described in detail below. Efficacy may be determined based on the ability of the antigen to provide at least about a 50% reduction in bacterial numbers and/or at least about a 40% decrease in mortality following experimental infection. Suitable experimental animals include mice, guinea pigs and primates.

Antigens having superior diagnostic properties may generally be identified based on the ability to elicit a response in an intradermal skin test performed on an individual with active tuberculosis, but not in a test performed on an individual who is not infected with *M. tuberculosis*. Skin tests may generally be performed as described below, with a response of at least 5 mm induration considered positive.

Immunogenic portions of the antigens described herein may be prepared and identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3d ed., Raven Press, 1993, pp. 243-247 and references cited

therein. Such techniques include screening polypeptide portions of the native antigen for immunogenic properties. The representative proliferation and cytokine production assays described herein may generally be employed in these screens. An immunogenic portion of a polypeptide is a portion that, within such representative assays, generates an immune response (e.g., proliferation, interferon- $\gamma$  production and/or interleukin-12 production) that is substantially similar to that generated by the full length antigen. In other words, an immunogenic portion of an antigen may generate at least about 20%, and preferably about 100%, of the proliferation induced by the full length antigen in the model proliferation assay described herein. An immunogenic portion may also, or alternatively, stimulate the production of at least about 20%, and preferably about 100%, of the interferon- $\gamma$  and/or interleukin-12 induced by the full length antigen in the model assay described herein.

Portions and other variants of *M. tuberculosis* antigens may be generated by synthetic or recombinant means. Synthetic polypeptides having fewer than about 100 amino acids, and generally fewer than about 50 amino acids, may be generated using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. See Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division, Foster City, CA, and may be operated according to the manufacturer's instructions. Variants of a native antigen may generally be prepared using standard mutagenesis techniques, such as oligonucleotide-directed site-specific mutagenesis. Sections of the DNA sequence may also be removed using standard techniques to permit preparation of truncated polypeptides.

Recombinant polypeptides containing portions and/or variants of a native antigen may be readily prepared from a DNA sequence encoding the polypeptide using a variety of techniques well known to those of ordinary skill in the art. For example, supernatants from suitable host/vector systems which secrete recombinant

protein into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant protein.

Any of a variety of expression vectors known to those of ordinary skill in the art may be employed to express recombinant polypeptides of this invention. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast and higher eukaryotic cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. The DNA sequences expressed in this manner may encode naturally occurring antigens, portions of naturally occurring antigens, or other variants thereof.

In general, regardless of the method of preparation, the polypeptides disclosed herein are prepared in substantially pure form. Preferably, the polypeptides are at least about 80% pure, more preferably at least about 90% pure and most preferably at least about 99% pure. In certain preferred embodiments, described in detail below, the substantially pure polypeptides are incorporated into pharmaceutical compositions or vaccines for use in one or more of the methods disclosed herein.

In one embodiment, the subject invention discloses polypeptides comprising at least an immunogenic portion of an *M. tuberculosis* antigen (or a variant of such an antigen) that comprises one or more of the amino acid sequences encoded by (a) the DNA sequences of SEQ ID NO: 1-12, 83, 102-108, 125, 127-137, 139 and 140; (b) the complements of such DNA sequences, or (c) DNA sequences substantially homologous to a sequence of (a) or (b). In a related embodiment, the present invention provides polypeptides comprising at least an immunogenic portion of an *M. tuberculosis* antigen having an amino acid sequence selected from the group consisting of sequences provided in SEQ ID NO: 16-33, 109, 126, 138, 141, 142 and variants thereof.

The *M. tuberculosis* antigens provided herein include variants that are encoded by DNA sequences which are substantially homologous to one or more of the DNA sequences specifically recited herein. "Substantial homology," as used herein, refers to DNA sequences that are capable of hybridizing under moderately stringent conditions. Suitable moderately stringent conditions include prewashing in a solution of 5X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5X SSC, overnight or, in the case of cross-species homology at 45°C, 0.5X SSC; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS. Such hybridizing DNA sequences are also within the scope of this invention, as are nucleotide sequences that, due to code degeneracy, encode an immunogenic polypeptide that is encoded by a hybridizing DNA sequence.

In a related aspect, the present invention provides fusion proteins comprising a first and a second inventive polypeptide or, alternatively, a polypeptide of the present invention and a known *M. tuberculosis* antigen, such as the 38 kD antigen described in Andersen and Hansen, *Infect. Immun.* 57:2481-2488, 1989, (Genbank Accession No. M30046), or ESAT-6 previously identified in *M. bovis* (Accession No. U34848) and in *M. tuberculosis* (Sorensen et al., *Infect. Immun.* 63:1716-1717, 1995). Variants of such fusion proteins are also provided. The fusion proteins of the present invention may include a linker peptide between the first and second polypeptides.

A DNA sequence encoding a fusion protein of the present invention is constructed using known recombinant DNA techniques to assemble separate DNA sequences encoding the first and second polypeptides into an appropriate expression vector. The 3' end of a DNA sequence encoding the first polypeptide is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide so that the reading frames of the sequences are in phase to permit mRNA translation of the two DNA sequences into a single fusion protein that retains the biological activity of both the first and the second polypeptides.

A peptide linker sequence may be employed to separate the first and the second polypeptides by a distance sufficient to ensure that each polypeptide folds into

its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may be from 1 to about 50 amino acids in length. Peptide sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons require to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

In another aspect, the present invention provides methods for using one or more of the above polypeptides or fusion proteins (or DNA molecules encoding such polypeptides) to induce protective immunity against tuberculosis in a patient. As used herein, a "patient" refers to any warm-blooded animal, preferably a human. A patient may be afflicted with a disease, or may be free of detectable disease and/or infection. In other words, protective immunity may be induced to prevent or treat tuberculosis.

In this aspect, the polypeptide, fusion protein or DNA molecule is generally present within a pharmaceutical composition and/or a vaccine. Pharmaceutical compositions may comprise one or more polypeptides, each of which

may contain one or more of the above sequences (or variants thereof), and a physiologically acceptable carrier. Vaccines may comprise one or more of the above polypeptides and a non-specific immune response enhancer, such as an adjuvant or a liposome (into which the polypeptide is incorporated). Such pharmaceutical compositions and vaccines may also contain other *M. tuberculosis* antigens, either incorporated into a combination polypeptide or present within a separate polypeptide.

Alternatively, a vaccine may contain DNA encoding one or more polypeptides as described above, such that the polypeptide is generated *in situ*. In such vaccines, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacterial and viral expression systems. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface. In a preferred embodiment, the DNA may be introduced using a viral expression system (e.g., vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer et al., *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

In a related aspect, a DNA vaccine as described above may be administered simultaneously with or sequentially to either a polypeptide of the present invention or a known *M. tuberculosis* antigen, such as the 38 kD antigen described above. For example, administration of DNA encoding a polypeptide of the present invention, either "naked" or in a delivery system as described above, may be followed by administration of an antigen in order to enhance the protective immune effect of the vaccine.

Routes and frequency of administration, as well as dosage, will vary from individual to individual and may parallel those currently being employed in immunization using BCG. In general, the pharmaceutical compositions and vaccines may be administered by injection (e.g., intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (e.g., by aspiration) or orally. Between 1 and 3 doses may be administered for a 1-36 week period. Preferably, 3 doses are administered, at intervals of 3-4 months, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of polypeptide or DNA that, when administered as described above, is capable of raising an immune response in an immunized patient sufficient to protect the patient from *M. tuberculosis* infection for at least 1-2 years. In general, the amount of polypeptide present in a dose (or produced *in situ* by the DNA in a dose) ranges from about 1 pg to about 100 mg per kg of host, typically from about 10 pg to about 1 mg, and preferably from about 100 pg to about 1  $\mu$ g. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will vary depending on the mode of administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, lipids, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (e.g., polylactic galactide) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268 and 5,075,109.

Any of a variety of adjuvants may be employed in the vaccines of this invention to nonspecifically enhance the immune response. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a nonspecific stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis*. Suitable adjuvants are

commercially available as, for example, Freund's Incomplete Adjuvant and Freund's Complete Adjuvant (Difco Laboratories) and Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ). Other suitable adjuvants include alum, biodegradable microspheres, monophosphoryl lipid A and quil A.

5 In another aspect, this invention provides methods for using one or more of the polypeptides described above to diagnose tuberculosis using a skin test. As used herein, a "skin test" is any assay performed directly on a patient in which a delayed-type hypersensitivity (DTH) reaction (such as swelling, reddening or dermatitis) is measured following intradermal injection of one or more polypeptides as described above. Such  
10 injection may be achieved using any suitable device sufficient to contact the polypeptide or polypeptides with dermal cells of the patient, such as a tuberculin syringe or 1 mL syringe. Preferably, the reaction is measured at least 48 hours after injection, more preferably 48-72 hours.

The DTH reaction is a cell-mediated immune response, which is greater  
15 in patients that have been exposed previously to the test antigen (*i.e.*, the immunogenic portion of the polypeptide employed, or a variant thereof). The response may be measured visually, using a ruler. In general, a response that is greater than about 0.5 cm in diameter, preferably greater than about 1.0 cm in diameter, is a positive response, indicative of tuberculosis infection, which may or may not be manifested as an active  
20 disease.

The polypeptides of this invention are preferably formulated, for use in a skin test, as pharmaceutical compositions containing a polypeptide and a physiologically acceptable carrier, as described above. Such compositions typically contain one or more of the above polypeptides in an amount ranging from about 1  $\mu$ g to  
25 about 100  $\mu$ g, preferably from about 10  $\mu$ g to about 50  $\mu$ g in a volume of 0.1 mL. Preferably, the carrier employed in such pharmaceutical compositions is a saline solution with appropriate preservatives, such as phenol and/or Tween 80<sup>TM</sup>.

In a preferred embodiment, a polypeptide employed in a skin test is of sufficient size such that it remains at the site of injection for the duration of the reaction  
30 period. In general, a polypeptide that is at least 9 amino acids in length is sufficient.

The polypeptide is also preferably broken down by macrophages within hours of injection to allow presentation to T-cells. Such polypeptides may contain repeats of one or more of the above sequences and/or other immunogenic or non-immunogenic sequences.

5

The following Examples are offered by way of illustration and not by way of limitation.

#### EXAMPLE 1

##### PURIFICATION AND CHARACTERIZATION OF *M. TUBERCULOSIS* POLYPEPTIDES USING CD4+ T CELL LINES GENERATED FROM HUMAN PBMC

*M. tuberculosis* antigens of the present invention were isolated by expression cloning of cDNA libraries of *M. tuberculosis* strains H37Rv and Erdman essentially as described by Sanderson et al. (*J. Exp. Med.*, 1995, 182:1751-1757) and

Two CD4+ T cell lines, referred to as DC-4 and DC-5, were generated against dendritic cells infected with *M. tuberculosis*. Specifically, dendritic cells were prepared from adherent PBMC from a single donor and subsequently infected with tuberculosis. Lymphocytes from the same donor were cultured under limiting dilution conditions with the infected dendritic cells to generate the CD4+ T cell lines DC-4 and DC-5. These cell lines were shown to react with crude soluble proteins from *M. tuberculosis* but not with Tb38-1. Limiting dilution conditions were employed to obtain a third CD4+ T cell line, referred to as DC-6, which was shown to react with both crude soluble proteins and Tb38-1.

Genomic DNA was isolated from the *M. tuberculosis* strains H37Rv and Erdman and used to construct expression libraries in the vector pBSK(-) using the Lambda ZAP expression system (Stratagene, La Jolla, CA). These libraries were transformed into *E. coli*, pools of induced *E. coli* cultures were incubated with dendritic cells, and the ability of the resulting incubated dendritic cells to stimulate cell

proliferation and IFN- $\gamma$  production in the CD4+ T cell line DC-6 was examined as described below in Example 2. Positive pools were fractionated and re-tested until pure *M. tuberculosis* clones were obtained. Nineteen clones were isolated, of which nine were found to contain the previously identified *M. tuberculosis* antigens TbH-9 and Tb38-1, disclosed in U.S. Patent Application No. 08/533,634. The determined cDNA sequences for the remaining ten clones (hereinafter referred to as Tb224, Tb636, Tb424, Tb436, Tb398, Tb508, Tb441, Tb475, Tb488 and Tb465) are provided in SEQ ID No: 1-10, respectively. The corresponding predicted amino acid sequences for Tb224 and Tb636 are provided in SEQ ID NO: 13 and 14, respectively. The open reading frames for these two antigens were found to show some homology to TbH-9, described above. Tb224 and Tb636 were also found to be overlapping clones.

Tb424, Tb436, Tb398, Tb508, Tb441, Tb475, Tb488 and Tb465 were each found to contain two small open reading frames (referred to as ORF-1 and ORF-2) or truncated forms thereof, with minor variations in ORF-1 and ORF-2 being found for each clone. The predicted amino acid sequences of ORF-1 and ORF-2 for Tb424, Tb436, Tb398, Tb508, Tb441, Tb475, Tb488 and Tb465 are provided in SEQ ID NO: 16 and 17, 18 and 19, 20 and 21, 22 and 23, 24 and 25, 26 and 27, 28 and 29, and 30 and 31, respectively. In addition, clones Tb424 and Tb436 were found to contain a third apparent open reading frame, referred to as ORF-U. The predicted amino acid sequences of ORF-U for Tb424 and Tb436 are provided in SEQ ID NO: 32 and 33, respectively. Tb424 and Tb436 were found to be either overlapping clones or recently duplicated/transposed copies. Similarly Tb398, Tb508 and Tb465 were found to be either overlapping clones or recently duplicated/transposed copies, as were Tb475 and Tb488.

These sequences were compared with known sequences in the gene bank using the BLASTN system. No homologies to the antigens Tb224 and Tb431 were found. Tb636 was found to be 100% identical to a cosmid previously identified in *M. tuberculosis*. Similarly, Tb508, Tb488, Tb398, Tb424, Tb436, Tb441, Tb465 and Tb475 were found to show homology to known *M. tuberculosis* cosmids. In addition, Tb488 was found to have 100% homology to *M. tuberculosis* topoisomerase I.

Seventeen overlapping peptides to the open reading frame ORF-1 (referred to as 1-1 - 1-17; SEQ ID NO: 34-50, respectively) and thirty overlapping peptides to the open reading frame ORF-2 (referred to as 2-1 - 2-30, SEQ ID NO: 51-80) were synthesized using the procedure described below in Example 3.

5           The ability of the synthetic peptides, and of recombinant ORF-1 and ORF-2, to induce T cell proliferation and IFN- $\gamma$  production in PBMC from PPD-positive donors was assayed as described below in Example 2. Figs. 1A-B and 2A-B illustrate stimulation of T cell proliferation and IFN- $\gamma$  by recombinant ORF-2 and the synthetic peptides 2-1 - 2-16 for two donors, referred to as D7 and D160, respectively.

10          Recombinant ORF-2 (referred to as M11) stimulated T cell proliferation and IFN- $\gamma$  production in PBMC from both donors. The amount of PBMC stimulation seen with the individual synthetic peptides varied with each donor, indicating that each donor recognizes different epitopes on ORF-2. The proteins encoded by ORF-1, ORF-2 and ORF-11 were subsequently named MTS, M11 and MSF, respectively.

15           Eighteen overlapping peptides to the sequence of MSF (referred to as MSF-1 - MSF-18; SEQ ID NO: 84-101, respectively) were synthesized and their ability to stimulate T cell proliferation and IFN- $\gamma$  production in a CD4<sup>+</sup> T cell line generated against *M. tuberculosis* culture filtrate was examined as described below. The peptides referred to as MSF-12 and MSF-13 (SEQ ID NO: 95 and 96, respectively) were found

20          to show the highest levels of reactivity. Two overlapping peptides (SEQ ID NO: 81 and 82) to the open reading frame of Tb224 were synthesized and shown to induce T cell proliferation and IFN- $\gamma$  production in PBMC from PPD-positive donors.

Two CD4<sup>+</sup> T cell lines from different donors were generated against *M. tuberculosis* infected dendritic cells using the above methodology. Screening of the *M. tuberculosis* cDNA expression library described above using this cell line, resulted in

25          the isolation of two clones referred to as Tb867 and Tb391. The determined cDNA sequence for Tb867 (SEQ ID NO: 102) was found to be identical to the previously isolated *M. tuberculosis* cosmid SCY22G10, with the candidate reactive open reading frame encoding a 750 amino acid *M. tuberculosis* protein kinase. Comparison of the

determined cDNA sequence for Tb391 (SEQ ID NO: 103) with those in the gene bank revealed no significant homologies to known sequences.

In further studies, CD4+ T cell lines were generated against *M. tuberculosis* culture filtrate, essentially as outlined above, and used to screen the *M. tuberculosis* Erdman cDNA expression library described above. Five reactive clones, referred to as Tb431, Tb472, Tb470, Tb838 and Tb962 were isolated. The determined cDNA sequences for Tb431, Tb472, Tb470, and Tb838 are provided in SEQ ID NO: 11, 12, 104 and 105, respectively, with the determined cDNA sequences for Tb962 being provided in SEQ ID NO: 106 and 107. The corresponding predicted amino acid sequence for Tb431 is provided in SEQ ID NO: 15.

Subsequent studies led to the isolation of a full-length cDNA sequence for Tb472 (SEQ ID NO: 108). Overlapping peptides were synthesized and used to identify the reactive open reading frame. The predicted amino acid sequence for the protein encoded by Tb472 (referred to as MSL) is provided in SEQ ID NO: 109. Comparison of the sequences for Tb472 and MSL with those in the gene bank, as described above, revealed no homologies to known sequences. Fifteen overlapping peptides to the sequence of MSL (referred to as MSL-1 -- MSL-15; SEQ ID NO: 110-124, respectively) were synthesized and their ability to stimulate T cell proliferation and IFN- $\gamma$  production in a CD4+ T cell line generated against *M. tuberculosis* culture filtrate was examined as described below. The peptides referred to as MSL-10 (SEQ ID NO: 119) and MSL-11 (SEQ ID NO: 120) were found to show the highest level of reactivity.

Comparison of the determined cDNA sequence for Tb838 with those in the gene bank revealed identity to the previously isolated *M. tuberculosis* cosmid SCY07H7. Comparison of the determined cDNA sequences for the clone Tb962 with those in the gene bank revealed some homology to two previously identified *M. tuberculosis* cosmids, one encoding a portion of hactoferritin. However, recombinant hactoferritin was not found to be reactive with the T cell line used to isolate Tb962.

The clone Tb470, described above, was used to recover a full-length open reading (SEQ ID NO: 125) that showed homology with TbH9 and was found to encode a 40 kDa antigen, referred to as Mtb40. The determined amino acid sequence

for Mtb40 is provided in SEQ ID NO: 126. Similarly, subsequent studies led to the isolation of the full-length cDNA sequence for Tb431, provided in SEQ ID NO: 83, which was determined to contain an open reading frame encoding Mtb40. Tb470 and Tb431 were also found to contain a potential open reading frame encoding a U-ORF-like antigen.

Screening of an *M. tuberculosis* Erdman cDNA expression library with multiple CD4+ T cell lines generated against *M. tuberculosis* culture filtrate, resulted in the isolation of three clones, referred to as Tb366, Tb433 and Tb439. The determined cDNA sequences for Tb366, Tb433 and Tb439 are provided in SEQ ID NO: 127, 128 and 129, respectively. Comparison of these sequences with those in the gene bank revealed no significant homologies to Tb366. Tb433 was found to show some homology to the previously identified *M. tuberculosis* antigen MPT83. Tb439 was found to show 100% identity to the previously isolated *M. tuberculosis* cosmid SCY02B10.

A CD4+ T cell line was generated against *M. tuberculosis* PPD, essentially described above, and used to screen the above *M. tuberculosis* Erdman cDNA expression library. One reactive clone (referred to as Tb372) was isolated, with the determined cDNA sequences being provided in SEQ ID NO: 130 and 131. Comparison of these sequences with those in the gene bank revealed no significant homologies.

In further studies, screening of an *M. tuberculosis* cDNA expression library with a CD4+ T cell line generated against dendritic cells that had been infected with tuberculosis for 8 days, as described above, led to the isolation of two clones referred to as Tb390R5C6 and Tb390R2C11. The determined cDNA sequence for Tb390R5C6 is provided in SEQ ID NO: 132, with the determined cDNA sequences for Tb390R2C11 being provided in SEQ ID NO: 133 and 134. Tb390R5C6 was found to show 100% identity to a previously identified *M. tuberculosis* cosmid.

In subsequent studies, the methodology described above was used to screen an *M. tuberculosis* genomic DNA library prepared as follows. Genomic DNA from *M. tuberculosis* Erdman strain was randomly sheared to an average size of 2 kb,

and blunt ended with Klenow polymerase, followed by the addition of EcoRI adaptors. The insert was subsequently ligated into the Screen phage vector (Novagen, Madison, WI) and packaged *in vitro* using the PhageMaker extract (Novagen). The phage library (referred to as the Erd  $\lambda$ Screen library) was amplified and a portion was converted into  
5 a plasmid expression library by an autosubcloning mechanism using the *E. coli* strain BM25.8 (Novagen). Plasmid DNA was purified from BM25.8 cultures containing the pSCREEN recombinants and used to transform competent cells of the expressing host strain BL21(DE3)pLysS. Transformed cells were aliquoted into 96 well microtiter plates with each well containing a pool size of approximately 50 colonies. Replica  
10 plates of the 96 well plasmid library format were induced with IPTG to allow recombinant protein expression. Following induction, the plates were centrifuged to pellet the *E. coli* which was used directly in T cell expression cloning of a CD4+ T cell line prepared from a PPD-positive donor (donor 160) as described above. Pools containing *E. coli* expressing *M. tuberculosis* T cell antigens were subsequently broken  
15 down into individual colonies and reassayed in a similar fashion to identify positive hits.

Screening of the T cell line from donor 160 with one 96 well plate of the Erd  $\lambda$ Screen library provided a total of nine positive hits. Previous experiments on the screening of the pBSK library described above with T cells from donor 160 suggested  
20 that most or all of the positive clones would be TbH-9, Tb38-1 or MTI (disclosed in U.S. Patent Application No. 08/533,634) or variants thereof. However, Southern analysis revealed that only three wells hybridized with a mixed probe of TbH-9, Tb38-1 and MTI. Of the remaining six positive wells, two were found to be identical. The determined 5' cDNA sequences for two of the isolated clones (referred to as Y1-26C1  
25 and Y1-86C11) are provided in SEQ ID NO: 135 and 136, respectively. The full length cDNA sequence for the isolated clone referred to as hTcc#1 is provided in SEQ ID NO: 137, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 138. Comparison of the sequences of hTcc#1 to those in the gene bank as described above, revealed some homology to the previously isolated *M. tuberculosis*  
30 cosmid MTCY07H7B.06

## EXAMPLE 2

### INDUCTION OF T CELL PROLIFERATION AND INTERFERON- $\gamma$ PRODUCTION BY *M. TUBERCULOSIS* ANTIGENS

5

The ability of recombinant *M. tuberculosis* antigens to induce T cell proliferation and interferon- $\gamma$  production may be determined as follows.

Proteins may be induced by IPTG and purified by gel elution, as described in Skeiky et al. *J. Exp. Med.*, 1995, 181:1527-1537. The purified polypeptides are then screened for the ability to induce T-cell proliferation in PBMC preparations. The PBMCs from donors known to be PPD skin test positive and whose T-cells are known to proliferate in response to PPD, are cultured in medium comprising RPMI 1640 supplemented with 10% pooled human serum and 50  $\mu\text{g}/\text{ml}$  gentamicin. Purified polypeptides are added in duplicate at concentrations of 0.5 to 10  $\mu\text{g}/\text{ml}$ . After six days of culture in 96-well round-bottom plates in a volume of 200  $\mu\text{l}$ , 50  $\mu\text{l}$  of medium is removed from each well for determination of IFN- $\gamma$  levels, as described below. The plates are then pulsed with 1  $\mu\text{Ci}/\text{well}$  of tritiated thymidine for a further 18 hours, harvested and tritium uptake determined using a gas scintillation counter. Fractions that result in proliferation in both replicates three fold greater than the proliferation observed in cells cultured in medium alone are considered positive.

IFN- $\gamma$  is measured using an enzyme-linked immunosorbent assay (ELISA). ELISA plates are coated with a mouse monoclonal antibody directed to human IFN- $\gamma$  (PharMingen, San Diego, CA) in PBS for four hours at room temperature. Wells are then blocked with PBS containing 5% (W/V) non-fat dried milk for 1 hour at room temperature. The plates are washed six times in PBS/0.2% TWEEN-20 and samples diluted 1:2 in culture medium in the ELISA plates are incubated overnight at room temperature. The plates are again washed and a polyclonal rabbit anti-human IFN- $\gamma$  serum diluted 1:3000 in PBS/10% normal goat serum is added to each well. The plates are then incubated for two hours at room temperature, washed and horseradish peroxidase-coupled anti-rabbit IgG (Sigma Chemical Co., St. Louis, MO) is added at a

30

1:2000 dilution in PBS/5% non-fat dried milk. After a further two hour incubation at room temperature, the plates are washed and TMB substrate added. The reaction is stopped after 20 min with 1 N sulfuric acid. Optical density is determined at 450 nm using 570 nm as a reference wavelength. Fractions that result in both replicates giving an OD two fold greater than the mean OD from cells cultured in medium alone, plus 3 standard deviations, are considered positive.

EXAMPLE 3PURIFICATION AND CHARACTERIZATION OF *M. TUBERCULOSIS* POLYPEPTIDES USING  
CD4+ T CELL LINES GENERATED FROM A MOUSE *M. TUBERCULOSIS* MODEL

5                   Infection of C57BL/6 mice with *M. tuberculosis* results in the development of a progressive disease for approximately 2-3 weeks. The disease progression is then halted as a consequence of the emergence of a strong protective T cell-mediated immune response. This infection model was used to generate T cell lines capable of recognizing protective *M. tuberculosis* antigens.

10                   Specifically, spleen cells were obtained from C57BL/6 mice infected with *M. tuberculosis* for 28 days and used to raise specific anti-*M. tuberculosis* T cell lines as described above. The resulting CD4+ T cell lines, in conjunction with normal antigen presenting (spleen) cells from C57BL/6 mice were used to screen the *M. tuberculosis* Erd  $\lambda$ screen library described above. One of the reactive library pools,  
15                   which was found to be highly stimulatory of the T cells, was selected and the corresponding active clone (referred to as Y288C10) was isolated.

                  Sequencing of the clone Y288C10 revealed that it contains two potential genes, in tandem. The determined cDNA sequences for these two genes (referred to as mTCC#1 and mTCC#2) are provided in SEQ ID NO: 139 and 140, respectively, with  
20                   the corresponding predicted amino acid sequences being provided in SEQ ID NO: 141 and 142, respectively. Comparison of these sequences with those in the gene bank revealed identity to unknown sequences previously found within the *M. tuberculosis* cosmid MTY21C12. The predicted amino acid sequences of mTCC#1 and mTCC#2 were found to show some homology to previously identified members of the TbH9  
25                   protein family, discussed above.

#### EXAMPLE 4

##### SYNTHESIS OF SYNTHETIC POLYPEPTIDES

5 Polypeptides may be synthesized on a Millipore 9050 peptide synthesizer using Fmoc chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation or labeling of the peptide. Cleavage of the peptides from the solid support may be carried  
10 out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile  
15 (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray mass spectrometry and by amino acid analysis.

From the foregoing, it will be appreciated that, although specific  
20 embodiments of the invention have been described herein for the purpose of illustration, various modifications may be made without deviating from the spirit and scope of the invention.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANTS: Alderson, Mark  
Dillon, David C.  
Skelky, Yasir A.W.  
Campos-Neto, Antonio

(ii) TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF  
TUBERCULOSIS AND METHODS OF THEIR USE

(iii) NUMBER OF SEQUENCES: 144

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SEED AND BERRY  
(B) STREET: 6300 Columbia Center, 761 Fifth Ave.  
(C) CITY: Seattle  
(D) STATE: Washington  
(E) COUNTRY: US  
(F) ZIP: 98104

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE: 05-MAY-1998  
(C) CLASSIFICATION:

## (vii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Makl, David J.  
(B) REGISTRATION NUMBER: 31,392  
(C) REFERENCE/DOCKET NUMBER: E16121.440C1

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 206-622-4900  
(B) TELEFAX: 206-682-6031

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1866 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(i) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

GGCTTGGTG ACCACCAACT TCTTGGTGT CAACACCATC CGATCGUCC TCACAGAGGC      60
CGACTACCTG CGCATGTGGA TCGAGCCCGC CACCGTCATG AGCCACTATC AAGCCGTGCG      120
CGACGAAATG TGGTGTCTCC ATGAATATAGC CAGTTCCGGA AAGCCGTGCG CCAGTATCAC      180
CAGCGGTGCG CCGGGTCCAC CGCGCTCGAC CACTGSCAGT CGCAGCGCGT TGGTATCAAC      240
TAACCGTNNN GTANGTCGC CCATGCTCTC ACCAANTCAC ACCGCGCATC GCGCTGAGAA      300
CGCGCTGGGG AGCANCTGCA GCGGATTGTC GCGGTTGCTG CGCGCATCA TGGTTCGGCC      360
GGCGGAGACA NTGGGGGCTC CCTTGACGTC CGGATGACAC TTCCTGTGCA GCTGKCATGG      420
CTACAGCTCA CAGTGACTGC CCGCGAGTTC CCGGCCAGGT CGAGTTCAAA TTCCGTTGAA      480
TTCGCGGACA AAAGCAGCAG GTCAACCAAC CGCAGTCAGT CGAGGTCGCC AAACGTGAGC      540
CAATCGGTGA AATGCGCTGC TCGACTGACA CGGTCACAGC GTTAGCGCGA CAGCACCGGA      600
ATAGCTCAGG CCGGCTATAG AGTCTATAG AAACATTTGC TGAATAGATT AACCGCTGTC      660
TTGGCGTGGT CTGTATACCG CTGCGCTGTC GACCGGTTGG CTGAGTAGCT GACCACCATG      720
TAACCGATTC TCGCGAGTTC TCTACTAAGC CGAGACACCG CATTTGTTGG GCTGCACTGC      780
AAATCGGTCC GAGCATGTAG CACTGCTGTT ATCCCGGAGT AGCAACCGAC CGGGAACGAG      840
GGCTATCCCA GTGCGCTCTC CAGCGAGGCG GTTTGCGTTT CCGTTGCGCG ATAACTCCCG      900
AGTGATATTC GGGTTTATCA NATTCAGGCT TTCTTCGCA AGGTACCGGT GTTCGCTATA      960
TTGGATATTC TCGGACGAGT AATTACTAAA ACTTCAGTGG TTTAGATAAG CGCGCGGCAA      1020
TACTTCGCGG ATCTTCGCGA GCGCAACGAA TTTCGATGCT GAGTTTCGCT CGGCTTATCA      1080
AACATGATCG GAGATATGTA CAGATCGGCC TAGCTAGGTC TTTAGCGGAG GCGATTAGG      1140
ACACCGGAGA TTTGCTTTGC CTGCGACCCA TGAAGCGGCC CCGCTTCGAG CGCGAATGCG      1200
GTGAGTGTG GTGGGTTGCT AGACGCTTGA TTGCGCCAGC GCGAGGTGTA TGTGCGGCG      1260
CAGCAGCGCT CCGCGCGGTA GCGCCATGAG CAGGATATAT AGACTCTCTC GCGACAGATC      1320
TCATACCGAT CGAAGCGGAA GCGCAGGCAAT CGAGCTCGGA GACACTGCTC TGGATCGCG      1380
CGCTCTACAC GCGGCTTGGC GCATTGTGCG ACCCGAGTTC CGGAGGAGCA AATGTGKCA      1440
GACATGTGAG TCGACACAGA GTGACATGCG CGTCTTACG AACTCAAAAC TGACGATCTG      1500
CTTAGCATGA AAAAACTGT TGCATCGCGC CAAGCATGAC AGCGAGCTG TAGGCTTACG      1560
CGTGCAATGG AGAAGCAAGG NTATGATGAG AATCGACGAC CATTTGAGATA GCGGCGAGCG      1620
ATGACAGAGC CTTTCATCAT CGATCGACGC ATCAGTGCCA TTGACGCTT GTACGACCTT      1680
CTGGGGATTC GAATACCCAA CGAAGGGGCT ATCCTTACT CTTCACTAGA GTACTTGAAA      1740
AAGGCTCTGG AGGAGCTGCG ACAGGCTTTC CCGGTTAATG CTGCTTAGG TTGGCGCGCG      1800
GACAAATGCG CGCGCAAAA CCGCAACGAC GTGAATTTTT TCCAGGAAC TCGCAGACCTC      1860
GATGCTCAGC TCATCAGGCT GTTCCA

```

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2395 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(iii) MOLECULE TYPE: DNA (genomic)

(iv) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

GGCAGCGCGT GCGCGCGCAA TACACGAAA TGCACAGGA ACTGCAAGC GTGCTGCTG      60
CGGTGAGCGC AAGCTCGTGG CAGGCGCCCA GCGCGAGCG GTTCTGTGTC GCCCATCAAC      120
CGTTCCGGTA TTGGCTAACC CACGCTGCCA CGGTGCGCAC CGCAGCAGCG GCGCGCGGACN      180

```

```

AAACGCGGCGT CCGCGGGTAT AGGTCCGCAT TGGGCGCAT GCTTACGCTA GCGGAGTTGG 240
CGGCGCAACCA TCGCATGACG GCGGCTCTGG TGACACACAA CTTCCTCGGT GTCAACAGCA 300
TCCGATGTCG CTCAACGAG GCGGACTTAC TCGGCACTTG GATTCAGGCT GCCACCGTCA 360
TGAGCGCTCA TCAAGCCGTC GCGGACGAAA GCGTGGCGCG GACCCCTGAG ACCGCGGCGG 420
CGCGCGCAGT MYTHAGCAG GCGGCGAGCT CGGCGGCTAG CAGCAGCTTC CCGGACCGGA 480
CGAAATATAT CCTGACAGTA CTCAGGATTT TCTCGGAGCT GCTGCGCTAT CTGCTCTTGG 540
AGCTGCTGCT GGGGCGCGCT GCGGACCTCA TGCGCCAGGT GTTGGACCTGG TTGATCTGGT 600
TGTGTTCTGG TCCAGTCTTC AGCTTTCTGG CCTACCTGGT GCTGGACCGA CTGATCTATT 660
TGGACCGGTT CCGCCCGGTC AGGAGTCCGG TCGCTTTGCT TGTCTGGAG TTACCGCAAC 720
CGCTCAAAAC CGCGCACCGA CTGACCGCTG CAGCTACCTT GATTTTGGAT CATTTCACTC 780
CGACTCGGCT CCGCGGATAT GTGCGCCGAC AAATGCTGTA CAGCGCGCCA ACGGAATTCG 840
GTGATCGGAC GTGCGGCTT GTGCAACCGG CTGCTGCGA ATTGCGCAGG AGTGTCTGTC 900
ATCAATGCTC CCGGAGACTT GCGGACACTC GCGGCGCTTG CCGACATCGA GATGATCTCC 960
CGCGAGATAG CAGGATTTGC CAACATGCTG ATGCTGCGGG GCTTGACCTG ACCGAGCGGG 1020
GAAGCTCTGA AGGAGACCAA GGTCTGCTTT CAGGCTGGTG AAGTGGGCGG CAGGCTGAGC 1080
GAAGCTCTGA CCGCTGCTGA AGAGACCGGA GCGGAGCTGG ACTGCTGAG CCGCGGTCGG 1140
CAGCAGTTGG CCGACGCGCT GCGGCAATA CGCACCGAAA TCATGGGGG GTGCGCGGCG 1200
TCGAGCGGGA TAGTCAACAC GGTGCAAGCC ATGATGAGCC TGATGGGCGG TGACAGGACC 1260
ATCGGCAACG TGAAGGATGC GTGCGGATAT GTCGGCGGCA TGCGGCTCTT GGGGAGCAAT 1320
CTGAGCGGGA CGGTCACCGA TGCGGACCAA ATGCGCACTT GGGCGAGGCT TTGCTCAAC 1380
GCTCTCACT CAGGCTCGCT GTGTAACAGC GATCGCGGCT GTGCGGCTTC GCGGCGGAGC 1440
TTGCGCGGGA TTGCGGAGG CGAGGAGCAC GCGCTGCTCA GTTGTGATAG ACCCTAGGCG 1500
GTCAACCTGC AACGAGCGCA GGAATACGNG ACATCTGCCC GGAAGGTTGG CACACTGAGC 1560
GCGCAACTGA AGCAAGTCTC CAGCAGCTTC AAAGCGGTG ACCGCTTACC CAGCAATTTG 1620
CTGCAAAATG AGCAAGGCTC CAGGCTCTTC GCGGAGGCA GCGGAGGCTT GCGGCAAGGC 1680
GTGCAAGGAT TGCTGATCTA GTTCAAAAGG ATGGGCTCAG GGTCTCAAGA GCGCGCGGAG 1740
TTCTGCTTGG GATCAAGAGC GATGCGGAGC AAGCGCTCAA TGGCGGGCTT CAGGATTTCA 1800
CGCGAGATTT TTTGAGGAGA CGAGTTCAAG AAGGCGCGCC AGATTTTCTT GTGCGCGGAT 1860
GCTCATGCGG CGGCTACTTT CCGTCAGAGC GCGCTGAAAT CCGGCGGAGC CAGGCGGATG 1920
GATCAGGTCA ACCATATGCT CCGTGTTCGG GATTCGCGCT GACCGAATAC CGAAGTGAAG 1980
GATGCGGAGA TAGGCTGTCG GGGGCTTCGG ACTGCGCTGC GGGATATTCG CGACTACTAC 2040
AACAGCGATA TGAAATTCAT GCTCATTCGG ACGATCGTTA TCGTATTCTT GATTTCTGTC 2100
ATTCTGNTGC GCGCACTTGT GGTTCGATA TATCTGATAG GCTCGGCTGT GATTTCTTAC 2160
TTGTGCGCTT TAGGCTATAG AACTTTGCTT TTCCAAATGA TATGCGGCA GGAATTCAT 2220
TGGAGCTGCG CCGGACTGTC GTTCATATTA TTGCTTGCCA TCGGCGGTGA CTACACATG 2280
CTGCTGATTT CAGCGATCTG CGAGG

```

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (a) LENGTH: 1742 base pairs
- (b) TYPE: nucleic acid
- (c) STRANDEDNESS: double
- (d) TOPOLOGY: linear

## (ii) MOLECULAR TYPE: DNA (genomic)

## (vi) ORIGINAL SOURCE:

- (a) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

CGGCTCTGTT TCAAGCTCAT AAGTTCGCTG GCGCACTCG CCGCGCTGCT ATATGCGACC 60
AATAAGCGCT GTTCCATGGA TACCGGAGCT GCACGACGCT AGAGCGGATC AGCGGAGCTG 120
GTGCGGAACA CTACGCGGTC CACGCTCAGC CTTGCGCGCT TGCGGAGGAT CCGGCGGAGG 180

```

```

TTCTCATGGT COTTRACGCC TTCCACACTG GCGACGGTGC GGCBCCGCGC GACCCTTGA 240
CCACAGCTGT GTCTCCGGAC CCGCGCGCGC GGTUCCAACA CCGCAGGATT GAGATGGAAG 300
CGGATACCGG STCGCATGAC ATGACCCGAC GCTCGATAGT ACBGGGUCU GACACCCGCC 360
AGATCATCCT TGAGTTCGGC CAGCGCGCGG TCGGTUCCGA ACNCGCGCMG CGCGCTGAA 420
CGTGAGCGCA CGATCCGCTG CACGACUAGC AGACGTCGCG GATACCCAA GCGCTTGGG 480
GTCCGCGMAT CCGGACNACH OTCGATGCTG TTCAGGTGAC GGAATGCTG GACCCGTGG 540
TGCTCGAGAT CGACAGCTGC CTGACATGCG AGGCGGTGCG GGTGCTGGGG ACACGGGCT 600
TGCTTCACGG GCTTTCGTCG ACCAGAGCCA GCATCAGATG CGCGCGCTG GACAGGATGT 660
CAGCTTCGCT CGGCTTCAGC GTCCGCGAGC GCTCAGGCG CCACTCTTGG AGAGAGGCT 720
TGCTGGGATT AATTGGGAGA GGAAGACAGC ATGCTGTTGG TGACGACACA GCGCGAAGCT 780
CTGGGAGCTG CGCGCGGAAA CCTACAGGAT ATTGCGACGA CAATGAACG CCGGAACCG 840
GCCCGGCTG CTCCAACGAC CGGAGTATGT CCGCGAGCGG CGGATGAAGT ATTAGCGCT 900
ACCGGGCTC AGTTTGTCTG GCACGCGCGG ATGTACCAA CGGTGAGCGC CCAGGCTCGG 960
CGCATTCAGC AATGTTCTGT GAACACGCTG GTGCGCAATT CTGGCTTATA CCGGGCCACC 1020
GAGCGGCGCA ACCGAGCGCG TCGCGGCTGA AGGCGCTGCG ACAGACCTGC TGAAGAGAG 1080
GGGGAACATC CGGATTTCTG GGTTCAGGGG TTGCGCTAGC GCGCAGCGGA TTCAATATC 1140
GGCTCCATA ACAGCGAGCG ATCTAGGCAT TCGCTCTTAA GAGACAGCGC AACATGGCT 1200
CAGCTTTTAT GAGGATATCG CATCGCATGC GCGACATGCG GCGCGGTTTT GAGCTGACCG 1260
CGGAGAGGAT GAGGAGAGAG GCTCGCGGGA TGTGGGCTC CGCGCAAAAC ATTTCGGT 1320
CGGCTGAGAG TGGCATGAGC GAGCGGAGCT CATTAGACCTG CATGACCTAG ATGATCAGG 1380
CGTTTCGCAA CATCTGAC ATCTGTCAGG AGTGGCTGA TGGGTGCTT CGCGAGCGCA 1440
ACAAATAGA ACAGAGAGAG CAGGCTTCCC AGCAGATGCT GAGCAATTAG CCGCGAAGC 1500
CAGAGGTGAG TACGATTTCT CAGATTAGGA GAAGACCAAT ATGACGATTA ATTACAGT 1560
CGCGGAGCTG GAGCTCTCAT GCGGCATGAT CGCGCTGAG GCGGCTGCG TTGAGTACG 1620
GCATGACGCG ATCGTTCGTG ATGTTTGGC CCGGCTGAC TTTTGGGCG GCGCGGTT 1680
GGTGGCTTGC CAGGATTTCA TTACCCAGTT GCGCGGTAA CTTGAGTGA TGTACAGGA 1740
GG

```

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2916 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

GTGATTTCCG TTGCGCGCGC CCGCGAAGAC CACCAACTCC GCTGGGGTGG TCCGACAGGC 60
GGTTCGCTGG GTCACTGCGC CGAATCCCAA TGATTTGTTG CTCTGTGGGG TTCTTGGGCT 120
CGATTACTCC CAGCGAAGAG AGGACGATGG TCGTTTGGCT CGGTCACTGG TACTTGGCGA 180
CGGCGATGGC CGCGTTTCTT ACCTCGATGG CACGACAGCT GACCTTCGGC CCGCGGGGCA 240
CAGCGCTGAG CTCGCGCGGA GCTTGTATCC CACGCGCACA ATTCGCGGGG TTGGGTGACG 300
GCCCGGCGGT GTGCGCGAGT TTAGCGCGGG CGGAGCGGCT CGGAGGTTG TCGGTGCGCG 360
CAGGTTAGGC GTTCGCGGCT CCGGCTTTCG CGGAGAGAGC TGAGCGCGGG ACCTCGATGT 420
CGTCTATGAG GGAAGCGCTCC AGCTTGGGTC AGGAGGCGCT GCTTGAAGCG ATACGCTGCG 480
CGAGGACGCG GCGGCTTACA GCGGCTTTCG CTCACCGATA CGGCTTCCCG CACAGCTGGA 540
TTACCGGCTG TCGGTGCGGG GATAGCTTTT GATCGGCTG TGGCGCGCGC CCGGAATATC 600
TGACGATGAG GATGACGCG CCGGCTCGGT AAACGCGGCA CACGCTACTA TCAATGCGCA 660
CGGCGGCGCT TGATGCGAAA TTGACCGTCC CGACGGGCGT TTATCTGGGG CAGAGATTTC 720

```

TCCCTCAGCTT GGTGGTGGG CCGATAAATA GGCTGGTCAQ CCGGACTCTT CCGGCTGAAT 789  
 TCGATGCTCTT GGGCGGCGAC TCGACGCGCA GTATCTCGAG TGGCGCGCAA ACCCGGTCAA 849  
 ACGCTGTTAT TGTGGGGTTA CCGACAGTGA ATTTGCGGTG CCAACTGGTG AACACTTGGG 909  
 AACCGGATGGC ATCGAAATGA ACTTGTTCGG TTGCACTGAT CTACTCTCTT CCGACAGAGCC 969  
 GTTCTCTGGG TTTATCTGGG GAGGAGAGCA GATGTGTGTT COTGACACCA CAGCGCGAAG 1029  
 CCCTTGGCAG TCGCGCGCGC AACCTACAGG GTATTGGGAC GACAAATGAG GCCCGAGAAAG 1089  
 CGGCGCGCGG TGCTTCACAC ACCGAGAGTAT TGGCGCGAGG CCGCGATGAA GTATCAGAGC 1149  
 TGACGCTGGC TCACTTTCCT GCGCACGCGG AGATGTACCA AACGCTCAGC GCGCGAGCGA 1209  
 CGGCGATTCA CGAAATGTTT GTGAACAGCG TCGTGGCGAG TTCTGGCTCA TACGCGCGCA 1269  
 CCGAGCGCGC CAACGCGAGC GCTGCGGCTT GACCGGGCTC GACCGAACCT CCGTGAAGGAG 1329  
 AGGCGGAGCA TCGCGAGTTC TCGGATCAGG GGTTCGCGCA GCGCCAGGCT GATTCTAGTA 1389  
 TCGGCTTCCA TAACAGCAGA CGATCTAGGG ATTCAGTACT AAGGAGACAG CCAACTATGG 1449  
 CTCACGTTT ATGACGAGTC CGCATCGCAT GCGGAGCATG GCGGCGGCTT TCGAGGTGCA 1509  
 CCGCGAGAGC GTGAGAGAGC AGCTTCGCGC GATGTGGGCG TCCGCGCAA ACATTTCGG 1569  
 TCGCGGCTGG AGTGCGATGG CCGAGCGGAC CTGCTAGAC ACCATGACCT AGATGAAATCA 1629  
 GCGCTTTCGG AACATCTGTA ACATGCTGCA CCGGATGCTT GACGCGCTGG TTCCCGAGCG 1689  
 CACACACTAC GAACAGCAGC AGCAGGCTTC CCAGCAGATC CTGAGCAGCT AGCGCGCGAA 1749  
 GGCACAGCGG GGTACGCTTT CTCACATTAG GAGAGACACCA ATATGACAGT TAATTACGAG 1809  
 TCGGCGGAGC TCGAGCTTCA TGGCGCGATG ATCGCGGCTC AGGCGCGGCT CCGTTCAGCG 1869  
 GAGCATAGCG CCGATGTTTG TGATGTGTTG GCGCGGGGTG ACTTTTGGG GCGGCGCGGT 1929  
 TCGGTGCTTT CCGAGAGATT CATTACCGAG TTGCGCGCTA ACTTTTGGT GATCTAGAGG 1989  
 CAGGCGAAGG CCGTACCGCA GAGGTGCGAG GCTGCGCGCA ACAACATGGG GCGAACCGAG 2049  
 AGCGCGCTGG GCTTCAGCTG GCGCTTAAAC TGAACCTGAG TCGCGCGAGC ACACCGAGCA 2109  
 CGCGGTGTC TGCTGTGCTC TGCACTTAAC TAGCAGTGA CCGCTGAGAT AGCATAGGAT 2169  
 CAACAGATGA CCGGAGCGGA CATCACCGTC AACCTGAGC GCTTCAGAT GCTTCAGCGC 2229  
 CTACTGATA TCGCGCATGT TCGCGCTGAG TTACCTTGCC GCGCTTACT CTGACAGGAT 2289  
 TCGAATAGCT GCGTAAAGTA GCACCGCGGG ATGCGCGGTA TCGCGGAGCA GCGCATGTCT 2349  
 GTCAACGAGC CCGTACAGGA ACAGTTCCTT GCGCGGATGA AGGTGCTTTC CCGACCTGAT 2409  
 CTTGAATGCG TCGCTCTGCT GTCACTGCGC AAGTGTGCTT ACAGGCTCAT ACAGCAGGAG 2469  
 AACACGCTGC CCGGTTTCGG TGACATGCTT GACAAAGAT TCGGCTGCTT GTTGGCTGGG 2529  
 CCGAGCGCAG ACTGGGTGTC GCGGTACGCG GTTGGCAGAT ACATCACCGT CGATGAGCTG 2589  
 ACGGCTCTCG ATAGCGGCTT GATCGCGGCA CTGCTAATGG ACGGCTTGGG CTGGAATTCAC 2649  
 CAGCGCGAGC CAGCGCGGAT CAACGCGGTC AACGTGCCAA TGGAGGAGAT CTGCTGCGCA 2709  
 ATTGCGCAG AGGCAAGGAG CCGTGTGCTT GACACGCGGA TGCAATACGA CCGTGGAGCG 2769  
 GCGGCGATCC TTGGCGATCT GGTGAGCGAC GACCGCGGCC CCGCGGAGGC TCTGCGAGAT 2829  
 CCGTGGGCTT TTCTCG

(i) INFORMATION FOR SEQ ID NO:5:

(a) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 980 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AACATGCTCT ACCGAGTCTG TGACGGGCTG GTTCCGAGC CCAACACTA CAGCAGCAA 68  
 GAGCAGGCTT CCGAGCAGAT CTCAGCAGC TACGCTCAGC CGCTGCAGCA CAATCTTTT 128  
 ACAAGCGAGG GAGAACAGGT TCGATGACCA TCACTATCA GTTCCGCTGAT GTCAGGCTC 188

|            |            |            |            |            |            |     |
|------------|------------|------------|------------|------------|------------|-----|
| ACGGGCGCAT | GATCTGGCGT | CAGCGCGGTT | TGCTGGAGGC | CGAACATGCG | GCATCATCTC | 240 |
| GTGATGTGTT | GACCTCGAGT | GACTTTTGGG | GGGCGCGCGG | TTCGGCGGCG | TGCGAGCGGT | 300 |
| TGATTAACCA | ATTGGGCGGT | AACCTCCAGG | TGATCTACGA | ACAGGCCAAC | GGCCAGCGGC | 360 |
| AGAGGTCGCA | GGTCTCCGGC | AACAGCAGGG | CGCAACCGGA | CAGCGCGCTC | GGCTCCAGCT | 420 |
| GAGGCTGACR | CCAGGCCAAG | GGCAGGGACG | TGGGTACGA  | GTGAGGCTTC | CTCGGCTGAT | 480 |
| CTTTCGGGTC | GCAGTCLAG  | TGGTCAGTGC | TGGGTTCTTC | GTGCTTTCTT | GCTTGGCGGG | 540 |
| TTCTTTCGGT | CTGGTCAAGT | CTGCTCGGGC | TGGGTTGAGG | ACCTCGAGGC | CGAGGTAGCG | 600 |
| CGTCCCTTCG | ATCCATTCGT | CTGTTTCTTC | GGCGAGGACG | GCTCGGACCG | CGCGGATGAT | 660 |
| CGAGGCGCGG | TGGGGGAGAG | TGCCACGACG | GTGCGTTTCG | GTGCTGACCT | CTCGGTTGAG | 720 |
| CGGTCTCTCG | GGGTTTCTGG | ACCGATTTCG | GCCTCGACGC | TTCTTGGGGA | AGCGGATGAA | 780 |
| CGCTAGACAG | TGCTTGGGGG | CGGTGTCGAG | GTGCTCGGCG | ACCGGGGAGA | GTCTTCTCGT | 840 |
| CAGAGGTCG  | AGTACCGGAT | CATATTGGCG | AACAACTGAT | TGGGCTTGGG | GCTGCTGCTA | 900 |

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1905 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

|            |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| GCTGCGCGGA | TGTGGGCGTC | CGCGGAAAGC | ATTTCGGGTC | CGGCTGCGAG | TGGCATGGCT | 60   |
| GAGGCGACCT | CGCTAGACAC | CATGGGCGAG | ATGAATCAGG | CGTTTCCCAA | CATCGTGAAC | 120  |
| ATGCTGCGAG | GGGTGCGTGA | CGGGCTCGTT | CGCGAGCGCA | ACAACTACGA | CGAGCAAGAG | 180  |
| CAGGCTCTCC | AGCAGATCCT | CAGCAGCTAA | CTTACGCCGC | TGCAGCACAA | TACTTTTACA | 240  |
| AGCGAAGGAG | AACAGGTTGG | ATGACCATCA | ACTATCATTT | CGGTGATGTC | GACCTTCAGG | 300  |
| GGGCGATGAT | CGCGCTCAG  | CGCGGCTTC  | TGGAGCGCGA | GCATCAGGCC | ATCATTCCTG | 360  |
| ATGTTTGGAC | CGCGGCTGAC | TTTTGGGCGG | CGCGCGCTTC | GGCGCGCTGC | CAGGGTTCGA | 420  |
| TTACCGCAT  | GGCGGCTAAC | TTCCAGGTGA | TCTACGAACA | AGCCAAACCC | CACGGGCGGA | 480  |
| AGGTGCAAGC | TGCGGCAAGC | AACATGGCGC | AACCGACAGC | CGCGCTCGGC | TGCGGTGGGG | 540  |
| CGTGACACGA | CGCGAGGCGC | AGCGAGGTTG | TGTACAGTGG | AAGGTTCTCT | CGGTGATCCT | 600  |
| TGCGGTCGCA | GTATAGTTGG | TGATGCTGCG | GGGTTTGGTG | GTCTGCTGCT | TGGCGGTTTC | 660  |
| TTGCTGCTG  | GTATGCTGTC | CTCGGCTCGG | GGTGAAGACC | TGAGGCGCGA | GGTGAAGACC | 720  |
| TCTCTCGGAT | CATTCGCTCT | GTCTTTCGCG | GAGGACGCT  | CGAGCGAGCT | CGATGATCGA | 780  |
| GGCGCGGTCG | GGGAAGATGC | CGAGCAGCTC | GGTTGGGCGT | CGTACCTCTC | GGTTGAAGCG | 840  |
| TTCTTGGGGG | CGACGCTCTG | GGCGGAGGCG | ACTTACGCTC | ATTTCGCTGC | ACTTACGCTC | 900  |
| GGTGGCGCAC | GACTATGAGC | AGCAGACCTC | TTTTCGCGAG | GGCGCTCGAG | GGGCTTCGCG | 960  |
| GTCCGCGCGA | CAGCTGTTTT | CGGTAAGTGA | CGCGCGGCAA | TTCTATGATG | GTATCGCGCG | 1020 |
| CGCGGAAAGC | CGCAAGGGAG | TGGGTGTGGA | CGGTTTCTTC | AAATGACCGG | CGAATCGCGG | 1080 |
| GGCGAGCTCG | CAGATTTTGC | AGATTTCTTC | ATCAACGCTC | CTTACCGCGA | CAGCGGCGCG | 1140 |
| ATCCAGGAGT | CTCATATCTT | TTTTATTCTT | CGGATCTGCG | AATTCGTCGA | ACGCTGCTCT | 1200 |
| TTGCGCGCTC | CGCAATGAGA | AAGCGATCTC | TTACCGCGCG | ATTGGAAGGA | GGGTTTCGCG | 1260 |
| AGTTCGCGGA | CAGCAATGAT | GTCTCTCTCT | CGATAGAGAG | GGGCTCATCA | ATGACGATGT | 1320 |
| GGTTCGCGAC | TAGCTACGCA | ACTGCGCGCA | GTCTTCTGTC | TTGCGCGCGG | CGCGCGCGCG | 1380 |
| GTGGAAGAG  | GTACCGGCTT | GGCTCTCTCT | CATCTGCTCT | GTGACAAACG | AGCGGCGCGT | 1440 |
| GGTTCGCGGA | TGATGAGCG  | CGGTGACGCT | GATGGTGTGA | CATGGCGCTC | TCCAAATGCA | 1500 |
| CGTTGCTATC | GATGGCTGTC | TGATGAGTGT | ATTTCAGGTT | TGCGCGCTCG | CGGATTCGCA | 1560 |
| CGGTTGCTGC | TGCGCTAGCG | CGAGACCGCG | TCTGGTCTCT | GACTGGCTCG | GAGTGAAGCG | 1620 |

|            |             |            |            |            |            |      |
|------------|-------------|------------|------------|------------|------------|------|
| CGACAGTGGG | CCATCTGCTGA | GCATCTGGGT | TGGGACACAG | CTCAGCGATC | TTGACATTTG | 1580 |
| CACACACGCT | CGCGGCTGCT  | GGCGGTGCAT | GTGCACTGT  | CCAGATAGGG | GGCGCCGCTT | 1740 |
| CTGGCGGTGT | CGCTGACCGG  | TCATTTGACT | CGCTCTGGGA | GTTCGCTGTC | CGAGTCGGAC | 1800 |
| ATGCGCGGGT | GGAGCGGGGG  | TAATGCGCAT | CTTGGCGGGG | CGAGCGCGCT | TCGCGCTCGG | 1860 |
| ACTTNGCGGT | CGCGGACAG   | ACGTGGAACC | GTACTCGGAC | CAATT      |            | 1905 |

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2521 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(iii) MOLECULE TYPE: DNA (genomic)

(v) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

|            |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| CGGATGCGG  | TGTTGGTTGG | TATTCGCCAA | ACCTTGGGCG | TGCTCCCGGG | GGTATCCAGG | 60   |
| TCCCGGTCGA | CCATCAGCGC | TGGACTGTTT | CTCGGACTCG | ACCGTGAAGT | GGCGCCCGGA | 120  |
| TTCCGATTCG | TGCTGGGCGT | TCCAGCGGTC | TTCCGCTCGG | GGTGTGTTCT | GTTCGCGGAC | 180  |
| GGATTCGACG | CGGTAACTCG | GGCATGAGG  | GGTACTGGCC | GGGATTTGTT | GGTGGCGGCT | 240  |
| CTGATCGCGT | TGCTGCTCGG | TCTGACCGCG | GTGGGCGGCG | TGCTGCGGTT | TCTGTGTCGA | 300  |
| CACCAACATG | AUTGCTTGTG | CGGCTACCGG | GTGCTGCTCG | GGAGGAGGAT | GGTGTGTCGT | 360  |
| CTGGCTACCG | GGACGCTAGC | CGGACATGGA | CGGTATCTCT | GGTACGCGAT | GGCGGTTTGA | 420  |
| CGTCGAAACG | CGCGGCGCTG | CTGGCGGCGG | GGTGGGCGCT | CGACTCGAG  | GGGAAAGGCG | 480  |
| GGGACGAGGC | CACCGGGGTC | ATCGATCGAA | TTGGTGAGCT | GGGATCGGGA | GGGGTGGGCT | 540  |
| CTTCTCCCAT | GGTGGCGGTT | CGAGCGACCG | TGGAAACCGT | GGCGGAGGCG | CTGTGCTTGG | 600  |
| AGCGGCTCAT | CGATGACCGG | TTCTCCGAAG | TGGACTACGG | CGAATGAGCT | GGCAGAAAAA | 660  |
| TGGTGAGCTT | GGTTCGAGAG | CGTTTGTGCG | GGGTAGTCCA | GGGCAACCTT | AGCGCGGGCG | 720  |
| TGTTTCGCGG | CGGTGAGGTT | TTGGCGGAGG | TGCAGAGCTG | GTGTTCTGGA | CGGATTTTGA | 780  |
| TGCGGGGGAA | CACCAAGACC | GGATCGGCAC | TGGCGGTCGC | CGGCGAAAAA | CCGCGCGGCA | 840  |
| ATAGGGGGAC | GGTGGCTTCC | AATGCGGGTG | GTACCGAGCG | GACCACTCTG | AACCTCCGAT | 900  |
| CGTCGAGGCG | AAGCGGATCG | CGCGCGCGCG | GTACCGGCTA | AGCGGTACCA | GAACCGGACG | 960  |
| GTAAATCTTC | GGCAATGTGC | GGTCGCGAGC | TTACCGAGAC | GTGACGAGAG | AGGCGCGGCG | 1020 |
| ATTGATGTTA | TGGATGGTGC | GGGCTTCCCA | NCGCGGCGGT | CGGAATACGT | AGCGCAGCGG | 1080 |
| ATGCGCGGGA | CGTGTGGCGG | ACCTCCGAGT | AGCGCAGGTC | GGCACTGACT | CGCGGCTGTC | 1140 |
| CGGCTGTCAG | ATGTTGAAAG | TGTCGACCGG | CTTGCTCAGG | CGATAATGCT | CTTGGAATAG | 1200 |
| CTCGGGCTGA | AAGCTACCGA | ACAGCGCGTC | CGAGATGATG | AGGATGCGCT | CATAGTTCTT | 1260 |
| GTCCAGATAC | ACCGGCTCCA | TTCGTTGGTG | GGCGCGGTTG | TGGAGCGGGG | TGTTGAGGAT | 1320 |
| GAAATCGGAC | CACCGCGGCA | GGCTGTGAT  | CGGCTCGGTC | TGCACCCGGA | ACTGTGAGAT | 1380 |
| CAAGTTCCAG | GACCAATTCG | AGAACACCAT | CGAAGGGGGA | AGCCCATCAT | GTGCGCGGCG | 1440 |
| AACCGCATAT | AGAACTTCGG | CGCTGTTGTT | CGAATTTCTG | CGCGAGCGCG | CTGCGCGGAT | 1500 |
| TGAATATTC  | GCTGGAGTGA | TGCGCTGCTT | GGGTAGCTCA | GATCGGCGGA | ACTCGGTTGG | 1560 |
| CGATCGGTTG | ATAGGAGTGA | TGCAGCGAGT | CGACACCAAC | GATCGGCTAT | ACCCAGGTTT | 1620 |
| ACACGCGGTC | GGCGGAGGTC | TGCAGGGGGG | CAGGTTAGCG | CATAGAACCA | GGTGTGAGAT | 1680 |
| CGGCGGCGAG | GGACTTCGAG | CGCGCGGTTG | TGGCTATCGA | AACGAGCGCC | ATGAGCTGTC | 1740 |
| TGCGCACCGA | GTGCGGCTGT | AGGTAAAGTC | CGGAGCGGCG | CGTGTGCTGC | CGGCTGTCAG | 1800 |
| CGGTCTCGAT | GCTTTCGAGC | TTGCGGCGCG | CGGTCTATTC | GAGGATCAGC | AGCAATAGAA | 1860 |
| AACATGGAAT | GGCGAACAGT | ACCGGCTGCT | GCATTTCTCT | GGCGAGCGCT | GAGAGGAATC | 1920 |
| CGCGCAGCGG | ATGGCTGAGT | CGACCTCGAT | AGACACCATG | ACCCAGATAT | ATGACGCGTT | 1980 |
| TGCAACATTC | GTGACATATC | TGCACGGGCT | GCCTGACCGG | CTGGTTGCGG | ACGCAACAAA | 2040 |

```

NTACGAACAG CAAGAGCAGG CCTCCACGCA GATCTTCAGC AGCTGACCCG SCGCCAGCAG 2199
TCAGGAGGAC ACATGACCAT CAATATTCAG TTGCGGCGGG TCGAGGCTCA CGCCGCCATG 2180
ATCCGCGCTC AGGCGCGGTC CTTGGAGGCC GAGCATCAGG GCATCATTTG TGATGTGTTG 2220
ACCGCGAGTG ACTTTTGGGG CGGGGCCGGT TCGCGGCGCT GCCAGGGGTT CATTAACCGAG 2289
CTGGCGCGTA ACTTCCAGGT GATNTACGAG CAGGCGACAG CCCACGGGCA GAGGTTGCGG 2340
GCTGCCCGCA ACAACATGCG ACAACGCGAC AGGCGCGTGG CCTCCAGCTG GGCATAGAAG 2400
TGGCTTAAGG CCGCGCGCGT CAATTACCAAC GTGGCGCGAC AGCGGTTGCT GGTGGGCCAG 2460
TTTGTTATCT GAACCACTAA CTACTTGCAG CTGCTAAGT CGCGCGCTTG ATCGCGGTTG 2520
GGATGTTGCT GAATGAGGAA GATGCGCTCA ATGCGCTTGT TGGCGAAGGG ATTGAGGCCA 2580
TCGTGTTTGG TACTTTAGGG GATCAGTGGT GATTGTGGGA GTCGCTGCTG CGCGAGCAGG 2640
TGGCGCGACT GCGCGAGGAA CTGGCGCGGG TGGAAGCATT GTTGAAGGAT CGCGGCTTCT 2700
TCGCGCGGCT CTGCGCGTTC TTGAGCGCGC GCAGGGGCGG GCGGTGAGAG CGGATGAGAG 2760
TGTATCTGCA GTTGATGTTT GGAAGTTTCC GTTACCGGCT GAGCATATAG TCGCTGTGCC 2820
GGGAGGCTGC TGATTCGATC ACCTGACCGC GTTTTTCGCC CATTGCGCTG GACCGGTTGG 2880
TGGCGCATCC GAGCAGATTG ATGAAGCTCA CCACGCGTTG C

```

## (ii) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1704 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (iii) MOLECULE TYPE: DNA (genomic)

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

## (vi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

CGGATGCTC GTCAAGGAG TCGACCGTCA CACGGACTG ATGACAAAT TCGAGGCGA 60
CGCGCGCTCG GCGATCTTGG GAGCGCGGAA CGCGCTGAG CGTCCGGAAG AGCGCGCGGT 120
GGCGCGCGCG CGGCGCATAN CGGAGCGGCT GCGGACGAG ATGCGCGAGG TCGAGCGCGG 180
CATCGCGGTC CGCGCAGGCG ANATGCTGCG CGGCAATGTC GCGGCCAAGC AAGGATTGTA 240
ATACAGATGG GTCGCGCAGG CGGTCAACCA NGCGCGCGCA TTGTGGAAG TGGCGAATTC 300
AGAGCGCGGG CGATTGGGTC TGCGCGCGCT GCGTATGCT CAGCGCAATC AAGGATCTAT 360
TTGCGCTTGG GCACAGCGTC CGGAAGTGGG CGATGGAAG CGCGAAGAGC GCGGCTGAGG 420
CGCGCAAGGG GTTGGCGGCG CGCGTTGAGG CATTGCGAG TGCTGGCTTG AGCGCGGTTG 480
CGGCGCGGCT CGTGCAGGCG GCGTGGGTGG GGGGATTGAA GGTTCGCGCG GTTTGAGCGG 540
CCAGCAGCCC GCGCGCGAGC CGCGCGGTGC TCGCGCGCT CACCGCTCTC GAGCGCGCGG 600
CCGCGCGTGA AGGTTGACA CAGCGGTTTG GCGGATGCG GCGTATGCGT AAGGTTGCGG 660
GAGTGTGCTT TAACAATTC GCGCGCGCTG GATACGATG CAGGCGGAGC GTGATGCGCG 720
AAGCGCGGCG TACGCGAGCA CCAACTAGT TCGTTGATCG AGGATCGAAT TCGAGGAGAT 780
AAGGCGAGGA ATTGATATGA GTTTCGCTTT TATGACGAG CGCGACGCGA TCGCGGAGAT 840
GCGCGCGGCT TTGAGGTCG AGCGCGAGAC GGTGAGGAG GAGGCTGCTG GATGTGCGG 900
GTGCGCGCMA AACATTTTCC GTGCGGCTG GAGTGGCAT GCGGAGCGCA CCTCGATGA 960
CAGCATGTCG CAGATGAATC AGCGGTTTTC CAACATCTGT AACATGCTGC ACGGCTGAG 1020
TGACGCGGCG GTTTCGAGCG CCAACAGCTA CGAGAGGAG GAGCAGGAGT CCGAGGAGAT 1080
CTTTCGCGCG TGACGCGGCG CGAGGACTCA GAGAGGACCA TGACCATCAA TTATCAATTC 1140
GAGGAGTTTC AGCGCATGCG CGCGATGATC CGCGGTTTGG CGGAGTTGCT GAGAGCGGAG 1200
CATGAGGCGA TCATTTCTCA TGTGTTGAGC CGGATGACT TTTGGGCGCG CGCGGCTTTC 1260
CGCGGCTGCG AGGAGTTTCT TACCGAGTTG GCGCGGAGT TCCAGGTGAT TTAGGACAG 1320
GCCAACCGCG ACGGCGAGAA GGTGCGGCT GCGGCAACA ACATGCGACA GAGGAGAGG 1380
CGCGTGGGAT CTAGCTGGGC CTAACCGGCG TGTAAATTC GGTTCGCGCA GAGCGCGCGC 1440

```

|            |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| ATCAGCGTMO | ACTTTGGGGC | CGATACACCG | GGCATTTTNT | NGTGGGGAAC | ACTTGGCCCG | 1500 |
| CGTCAGTTC  | CGCGTTCCCG | TTGTGAGGCG | ACGTACTCGG | TGATGGCTTT | GACGACCGCT | 1550 |
| TGGCGCGGCG | GGCCAAATCA | TTGGTGGGCG | TTGGCTNTAG | CCCATTCGTC | CGACGCCGCG | 1600 |
| GGCGCGCGCA | GTATGCGCTT | GAATTAAGCA | ATCAGACGAC | GGGCGACGAG | CTCATAGGAG | 1650 |
| TGAAGGTTT  | CGTGGCGGCG | GGCC       |            |            |            | 1700 |

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2286 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

## (iv) SEQUENCE DESCRIPTION: SEQ ID NO:9:

|            |            |             |            |            |            |      |
|------------|------------|-------------|------------|------------|------------|------|
| CGCTCTTGGC | GTCTGGGGCG | ATTGTGATCT  | GGGGCAATTG | CGCTTCCAGC | CAGAGCGCGC | 60   |
| CCAGCTTTTC | GATCCAGGCG | GGGACCCGGA  | TTGGCAGCGC | CGGACCCGCG | ACGAGATTCT | 120  |
| CGCTGAATT  | CTGGCTCACT | TGGAGTTCG   | GGGGGTGATC | CTGTTGCGCA | NCAGGCTTTC | 180  |
| GAACGGGCT  | GAAGCGCTG  | CGGTAGCGCC  | AGCGGTATCG | CGGTCAAGCG | GACCGGATTC | 240  |
| CGGATGCTT  | TGCGGCGCGA | CGTGGAGCGC  | CGCGGCTCCA | CGAAGAGGCT | CACGGTGGAG | 300  |
| CAGGCAAGCA | GATGCAAGCG | GAGGATCACT  | CGGAAGTGCC | GAATTCGCGA | CGAGAGGTTC | 360  |
| TGGAATCTCA | GCATACGCGC | CGCGGCGCGA  | TCCTGTTGGA | CTAGACCATC | GGGACGAGTC | 420  |
| CGGACGCGCA | CGTGGCGGCT | TTGATCGAAA  | ACAGCGAGCG | GGTGGTGCGC | GTGACCGCGC | 480  |
| TGCTCTTCAC | TTTCTGCGAT | GATCACTGCG  | ATTCGCTGCT | GGACAGCTTC | TCCGAGCGTG | 540  |
| AGGCGGGGCT | GCTGGCGGTA | CGCTTGGCGC  | TTACGAGCGC | CGAGCGCGCG | ACCTTTGAGC | 600  |
| AGATUGCGCA | GGTTTACGGC | GTCAGCCGCG  | AAAGCATCGC | CCAGATCGAA | TCCAGACTCA | 660  |
| TGTCGAAGTT | GGCGCATCGG | AGCGGATCAC  | AGGTCTCGCG | CGACTATGCT | GGCGAATTCG | 720  |
| CGACGAGCGC | TTTTAGGTTC | CAGCGCCGGA  | CGGTGGAGGA | GGAGGCTCGC | CGGATGTGGG | 780  |
| CGTCGCGCGA | AAGCATTTTC | GGTGGCGGCT  | GGAGTGGCAT | GGCGGAGCGC | ACCTCGCTTC | 840  |
| ACGCAATGCG | CCAGATGATC | CAGCGCTTTC  | CGAACGTCCT | GAACATGCTC | CACCGGCTTC | 900  |
| GTGACGGGCT | GGTTCGCGAC | GGCAAGCACT  | ACGACAGCAG | AGAGCAGGCG | TCCGAGGAGA | 960  |
| TGCTCGAGCA | CTGAGCCGCG | CGGAGGACTC  | AGGAGGAGAC | ATGACCATCA | ACTATTCATT | 1020 |
| CGGGAGCGTC | GAGCTCTCAT | CGGCGCATGAT | CGGCGCTCTG | CGCGAGTTTC | TGGAGCGCGA | 1080 |
| CGGATCAGCG | ATCATTTCTG | ATGTGTTGAC  | CGGAGTTCAC | TTTTGGGCGG | GGCGCGGTTT | 1140 |
| GGGCGCTCGC | CGGCGGTCGA | TTACCGCATT  | GGCGCGTACG | TTCCAGGTGA | TCTACGAGCA | 1200 |
| GGCGAACCGC | CACGCGCAGA | AGGTGCGAGC  | TGCGCGCAAC | AACATGCGAC | AACCGGACAG | 1260 |
| CGCGCTCGCG | TCCAGCTGGG | CGTACCGCGC  | GTCTTAAGTT | GGTTCGCGCG | AGCGCGGCGC | 1320 |
| GATCAGGCTC | GACTTTGGCG | CGCGATACAC  | GGCGATGTCG | TGCTGGGGA  | CACCTGGCGC | 1380 |
| GGGTGAGGTC | CGCGCTTTCG | CTTGTTCGCG  | GACGTCTCGC | GTAATGCTCT | TGAGGAGCGC | 1440 |
| TTGCGCGCGC | GGCGCATCA  | ATTGCTCGCG  | CTTGGCTCTA | GGCTGCTCGC | GAATTCGCGA | 1500 |
| CGAGGCTGCT | GGTGGCGCGC | TATCGGCGAC  | AGTTCGAGTC | CACGACGAGC | TGATCTGAGT | 1560 |
| GCTGGGTTTC | GGGAGTTTCG | GATCGCGGCT  | GTGCGCGCGA | AGGCGCATCG | CGGCGCATAT | 1620 |
| GGGCGCTGAA | GGCGCTTCG  | AGTACAGCGT  | CTGCGGCGAC | CGGCTCAAGC | AGGCGGCGCG | 1680 |
| CTCAGCGCGA | CTGGCCAAAG | TGAGGATGCG  | CGAGTTTCCT | CGCTTGGCGA | TGCGGCTGAG | 1740 |
| TGGCGCGCTG | GAGCGCGGAG | CATTGTGTTG  | GGATGTTGGC | GAGGTGGTTG | AGCTCGCGCG | 1800 |
| AGGTGCTTCA | CGGACCTAAC | TAGGCGAGCG  | AATGAATTCG | CGGCGAGCGC | AAGAGGTTTC | 1860 |
| CAGCGAGTCA | CGCGGCTAGT | CGCGCTTTCG  | TGCTTTCTTC | CGCGGACCTT | TGCGGCGAGC | 1920 |
| TTTCTGCTCT | GGCGGTTTTG | CGGAGCGCGC  | GGGTGCGCGA | TGCGGCAACA | GCTTGGCGAG | 1980 |
| GGCTCTGCTG | GTATAGGAGG | CGAGTTTCTC  | GTCTTAAGCG | AGGCTGCGAT | TGATCTGAGC | 2040 |

|            |            |            |            |            |             |      |
|------------|------------|------------|------------|------------|-------------|------|
| GTGGGTGACG | TACGGCTCGA | ATCGGCCCTC | CTTGAATGAC | ATTGGCTTGC | CAGACGCGCG  | 2100 |
| ATTGNTTCCG | AGCTCGGCA  | CGCGCGAGAC | CGAAGCTCTT | TGCGCGCAC  | GACNTTTGCG  | 2150 |
| CTCTGNTGAG | ATTTTCAGGG | CTTCGTCGAG | CGGATGCTG  | AATATATGGT | CTTCCGTTGAC | 2200 |
| CAGTGATCGA | GATTCGTTGC | CGCGCTTTAG | ATACGCTGCG | TAGCGCCCTT | TCTGCGCGGT  | 2250 |
| GATNTC     |            |            |            |            |             | 2290 |

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1156 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

|             |            |            |            |            |            |      |
|-------------|------------|------------|------------|------------|------------|------|
| GGGCATCTTC  | CGGACCGCG  | CCTCGATCAT | CGGCTCTGTC | GGAGCGCTGC | TGCGCGAAGA | 60   |
| ACACGACGAA  | TGGATCGAAG | GACGGCGCTA | CGTGGGCTTC | GAAGTCTCTA | CGCGAGCCCG | 120  |
| AGCAGCAGTG  | ACCGACACCG | AAGACCGGCG | AGCGAGCGAA | CGACGACGAC | CGCGAGCTG  | 180  |
| ACACCTTAGA  | CTGCGACCGG | AAGGATCAGG | CGAGGAGCTT | TCAGTCTGAC | ACCGAGTCCC | 240  |
| TGGGCTTTGCG | CTGGTCTGCG | CGCGAGCTCG | AGCGGACCGG | CGTCTGGGTT | TGCGCGCTGT | 300  |
| TGTTGCGCGG  | AGCTGCGAGC | TTCTGCTCGT | GGGCTTTGCG | CTGCTGCTAG | ATCGAGTGGG | 360  |
| AGTTACCGGG  | CAACTGGGTA | ATGAAGCCCT | GGCAGGCTGC | CGAGCGGCGG | CGCGCCCAAA | 420  |
| AGTACTGCTG  | GGTCAACACA | TGCGGAATGA | TGGCTGATG  | CTCGGCTCTC | AGCGACCGCG | 480  |
| CCTGAGCCCG  | GATCATGGCG | CGGTGAGCGT | CGACATCGCC | GAGCTGATAG | TTGATGCTCA | 540  |
| TGGAAGCTGT  | TCTCTTTGCG | TTGTAAAGAT | ATTGTGCTGC | AGCGGCTGAC | GTTAGCTGCT | 600  |
| GAGGATCTGC  | TGGGAGGCGT | GTCTTTGCTT | CGTGGCGGAT | TGGGACGAG  | AGGCGGCTTT | 660  |
| CGAAGGAATC  | CTTTGAGGAT | TGCGCAAGCG | CGTGCAGCGA | GCATGGGGTG | AGCTCGCGAG | 720  |
| CGCGCCCGCG  | TGGCAACGCT | TGCGGCTCGA | GAGGAGCTG  | GAGGATACG  | AGTGACAAAC | 780  |
| GAGCTCCCGG  | AGGTCGAGGA | CGGTGACGCG | GGTCCAGCTC | CGGCTCTCTC | TGCTGCGCGG | 840  |
| CGACCTTTGT  | CGAGCTCTG  | GGTTTACGAG | GGCGCGGGGT | ACGACCTGAG | TGAGTGGATT | 900  |
| TGCAAGGATC  | CGGCGCGGCG | CTTTTTCATT | GGGCGGACCA | AGAACCGCGA | CATCACCGCA | 960  |
| AGTGTCAGAT  | CTTACCATCG | TGATCGCGCG | ATTGTCGAGC | GATTCCTGCA | CGCGGAGGAT | 1020 |
| GCTTGGGCGG  | CGGACGCGAC | CGCTAGGAGC | ATCCACCGCA | AGCAGATGCG | ACCGGCTATT | 1080 |
| CTGTTCAAGG  | ACGACTTCAG | CGCTGGCGGG | GACACCGCGA | ATATCTGATT | NGAGA      | 1136 |

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 967 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

|             |             |             |            |             |            |     |
|-------------|-------------|-------------|------------|-------------|------------|-----|
| TGAGCGCCAA  | CGCTACCGTC  | GGTTGCTGAC  | AGGAGACGGA | TGGCTTCTCT  | CGGCGACTGC | 60  |
| CGCTAGGCTC  | GGGGATCACT  | CGGCGTAAACG | GGCGCTTTGC | CCACCGATAT  | GGCTTCCGTC | 120 |
| ACAGTGTGGT  | TGCGCGCGCG  | CCATCGGCGG  | GATACAGGCA | TGAGCTCAGC  | TGGGAGGAAA | 180 |
| TGACAAATGCT | CCGAAAGGCG  | TGAGCACTCG  | AAGACAATTA | AGCAGGAGAT  | CGTATCGGCT | 240 |
| TTGTGACTAC  | CCAACACGAA  | GCATTCGCGG  | CGGCGGACGG | CGCTTCTGAG  | GGAAATCGCT | 300 |
| CGGATTTGAA  | CGCGGAGAAAT | GGGCTTGGGG  | CGACTCCGAC | GACGCGGCTG  | GTGCGGCGCG | 360 |
| CGCGGATGAA  | NTGTGCGGCG  | TGACGCGGCG  | TGAGTTGCGG | GCACGACCGCC | AGATTTATCA | 420 |
| GGCGGTCAGC  | CGCGGAGCGG  | CGCGGATTTA  | CGAGATGPTC | GTACACACTC  | TACAGATGAG | 480 |
| CTCAGGCTG   | TATGCTGCTA  | CGGAGCGGCG  | CAATCGGCGG | GGGCGCGGAT  | AGAGGAGTCA | 540 |
| CTGCGATGGA  | TTTTTGGGCG  | TTCGCGCGCG  | AGTTCGATTC | GATGCGGATG  | TATCGGCTTC | 600 |
| CTGGCTCGGC  | ACCAATGGTC  | GGTGGCGGCT  | CGGCTTGGAA | CGGTTTGGCG  | GGGAGCTGGA | 660 |
| GTTCGCGGCG  | CACCGGTTAT  | GAGCGGCTGA  | TGACTGAGCT | CAGCAGTGAG  | GGGTGGCTAG | 720 |
| GTTCGCGGCT  | AGCGGCGATG  | CGCGGAGCGG  | TTGCGGCTTA | TTTGGCTTGG  | ATGAGTGGCG | 780 |
| CTGCGCGGCG  | AGCGGAGCGG  | CGCGGCGGCG  | AGCGGAGCGG | CGCGGCGGCG  | CTTTTGGAGG | 840 |
| CGGCTTTTTC  | CGCGGAGGCG  | CTTCGCGGCT  | TGATCGGCGG | CAACCGGCGG  | TGTTGATGCG | 900 |
| AGCTGATGTC  | GAGGAATGTC  | TTTGGTCAGA  | AGACTCGGCG | GATCGGCGCG  | GGGGAAGCTC | 960 |
| AGTACGG     |             |             |            |             |            | 967 |

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 585 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

|            |            |            |            |            |             |     |
|------------|------------|------------|------------|------------|-------------|-----|
| TGGATTCGGA | TAGCGTTTC  | GACCGCTCGA | CGGCGACCA  | CGGCGCGAG  | GGCTCGGAAC  | 60  |
| GGGCGCGCGG | GAGCGTGGGA | TTGCGTCGGA | CGGCAACGAA | AGAACCGCGG | GTCCGCGCGG  | 120 |
| TGGCGCTGAC | CGGACTGGCC | GGTGAATGAG | TGGCGAAGCG | CGGCGGATG  | CCGATGGTGC  | 180 |
| CGGCGGACTG | GGAGGAGCGG | AGCAACGAGC | CGGAGGCGCG | CGAGGATGCG | GCGAGAGGGG  | 240 |
| GGGCGGAGCG | CTTACGCGAG | GACAGCGAGT | AACCGAATTC | CGAATCAAGT | GGACCGGTAC  | 300 |
| GGGTGCAAGG | GAGAGATGTT | ATGAGGCTTT | TGAGTGTGCA | TATCCGACAG | TTGCTGGCGCT | 360 |
| CGGATCGGCG | GTTCGCGCGC | AAGCGCGGCG | TGATGGGCGA | CAGGATCGGT | GAGCGCGGAC  | 420 |
| AGGCGCGGAT | GTGCGGCTAG | GGCTTTACCC | AGGCGGAGTC | GTGCGGCGCG | TTTGGGCGCG  | 480 |
| CGGATCGGCG | GTTCGCGCGG | GGGCGCGCGA | AGTTCACAC  | CTTGTGGGAT | GTGCGGAGCG  | 540 |
| CGAATCGGCG | TGAGCGCGCG | GGTACCTATG | TGGCGCGCGA | TGCTG      |             | 585 |

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

Ala Leu Val Thr Thr Asn Phe Phe Gly Val Asn Thr Ile Pro Ile Ala
 1           5           10           15
Leu Asn Glu Ala Asp Tyr Leu Arg Met Trp Ile Gln Ala Ala Thr Val
 20           25           30
Met Ser His Tyr Gln Ala Val Ala His Glu Ile Trp Cys Leu His Glu
 35           40           45
Xaa Ala Ser Ser Gly Lys Pro Trp Ala Ser Ile Thr Thr Gly Ala Pro
 50           55           60
Gly Ser Pro Ala Ser Thr Thr Arg Ser Arg Thr Pro Leu Val Ser Thr
 65           70           75           80
Asn Arg Xaa Val Xaa Ala Pro Ile Val Ser Pro Asn His Thr Gly His
 85           90           95
Arg Pro Glu Lys Gly Leu Gly Ser Xaa Gln Arg Arg Leu Ser Arg Val
100           105           110
Leu Pro Arg Ile Ile Asp Arg Pro Ala Gly Pro Xaa Gly Pro Pro Leu
115           120           125
Thr Ser Gly Ser His Phe Leu Cys Ser Trp His Gly Tyr Ser Ser Gln
130           135           140

```

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 352 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

His Ala Leu Ala Ala Gln Tyr Thr Glu Ile Ala Thr Glu Leu Ala Ser
 1           5           10           15
Val Leu Ala Ala Val Gln Ala Ser Ser Trp Gln Gly Pro Ser Ala Asp
 20           25           30
Arg Phe Val Val Ala His Gln Pro Phe Arg Tyr Trp Leu Thr His Ala
 35           40           45
Ala Thr Val Ala Thr Ala Ala Ala Ala His Xaa Thr Ala Ala Ala
 50           55           60
Gly Tyr Thr Ser Ala Leu Gly Gly Met Pro Thr Leu Ala Glu Leu Ala
 65           70           75           80
Ala Asn His Ala Met His Gly Ala Leu Val Thr Thr Asn Phe Phe Gly
 85           90           95
Val Asn Thr Ile Pro Ile Ala Leu Asn Glu Ala Asp Tyr Leu Arg Met
100           105           110
Trp Ile Gln Ala Ala Thr Val Met Ser His Tyr Gln Ala Val His His
115           120           125

```

```

Glu Ser Val Ala Ala Thr Pro Ser Thr Pro Pro Ala Pro Gln Ile Val
130 135 140
Thr Ser Ala Ala Ser Ser Ala Ala Ser Ser Ser Phe Pro Asp Pro Thr
145 150 155 160
Lys Leu Ile Leu Gln Leu Leu Lys Asp Phe Leu Gln Leu Leu Arg Tyr
165 170 175
Leu Ala Val Gln Leu Leu Pro Gly Pro Leu Gly Asp Leu Ile Ala Gln
180 185 190
Val Leu Asp Tyr Phe Ile Ser Phe Val Ser Gly Pro Val Phe Thr Phe
195 200 205
Leu Ala Tyr Leu Val Leu Asp Pro Leu Ile Tyr Phe Gly Pro Phe Ala
210 215 220
Pro Leu Thr Ser Pro Val Leu Leu Pro Ala Val Gln Leu Arg Asn Arg
225 230 235 240
Leu Lys Thr Ala Thr Gly Leu Thr Leu Pro Pro Thr Val Ile Phe Asp
245 250 255
His Pro Thr Pro Thr Ala Val Ala Gln Tyr Val Ala Gln Gln Met Ser
260 265 270
Gly Ser Arg Pro Thr Gln Ser Gly Asp Pro Thr Ser Gln Val Val Gln
275 280 285
Pro Ala Arg Ala Gln Phe Gly Thr Ser Ala Val His Gln Ile Pro Pro
290 295 300
Arg Pro Ala Asp Thr Arg Arg Ala Cys Arg His Arg Asp Asp Val Pro
305 310 315 320
Arg Asp Ser Arg Ile Ala Gln His Arg Asp Gly Ala Gly Leu Asp Pro
325 330 335
Thr Gln Arg Gly Thr Ser Gln Gly Asp Gln Gly Leu Val Ser Gly Trp
340 345 350

```

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 141 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```

Met Asp Phe Gly Ala Leu Pro Pro Gln Val Asn Ser Val Arg Met Tyr
1 5 10 15
Ala Val Pro Gly Ser Ala Pro Met Val Ala Ala Ala Ser Ala Tyr Asp
20 25 30
Gly Leu Ala Ala Gln Leu Ser Ser Ala Ala Thr Gly Tyr Gln Thr Val
35 40 45
Ile Thr Gln Leu Ser Ser Gln Gly Trp Leu Gly Pro Ala Ser Ala Ala
50 55 60
Met Ala Gln Ala Val Ala Pro Tyr Val Ala Trp Met Ser Ala Ala Ala
65 70 75 80
Ala Gln Ala Gln Gln Ala Ala Thr Gln Ala Arg Ala Ala Ala Ala

```

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     | 95  |     | 90  |     | 95  |     |     |     |     |     |     |     |     |     |     |
| Phe | Glu | Ala | Ala | Phe | Ala | Ala | Thr | Val | Pro | Pro | Pro | Leu | Ile | Ala | Ala |
|     |     | 100 |     |     |     |     |     | 105 |     |     |     |     |     | 110 |     |
| Asn | Arg | Ala | Ser | Leu | Met | Gln | Leu | Ile | Ser | Thr | Asn | Val | Phe | Gly | Gln |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     |     | 125 |     |     |
| Asn | Thr | Ser | Ala | Ile | Ala | Ala | Ala | Glu | Ala | Gln | Tyr | Gly |     |     |     |
|     |     | 130 |     |     |     | 135 |     |     |     |     |     | 140 |     |     |     |

## (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 58 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Ser | Arg | Phe | Met | Thr | Asp | Pro | His | Ala | Met | Arg | Asp | Met | Ala |
| 1   |     |     | 5   |     |     |     | 10  |     |     |     |     |     | 15  |     |     |
| Gly | Arg | Phe | Glu | Val | Sis | Ala | Gln | Thr | Val | Glu | Asp | Glu | Ala | Arg | Arg |
|     |     |     | 20  |     |     |     | 25  |     |     |     |     | 30  |     |     |     |
| Met | Trp | Ala | Ser | Ala | Gln | Asn | Ile | Ser | Gly | Ala | Gly | Trp | Ser | Gly | Met |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |
| Ala | Glu | Ala | Thr | Ser | Leu | Asp | Thr | Met | Thr |     |     |     |     |     |     |
|     |     | 50  |     |     |     | 55  |     |     |     |     |     |     |     |     |     |

## (2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 67 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Thr | Ile | Asn | Tyr | Gln | Phe | Gly | Asp | Val | Asp | Ala | His | Gly | Ala | Met |
| 1   |     |     | 5   |     |     |     | 10  |     |     |     |     |     | 15  |     |     |
| Ile | Arg | Ala | Gln | Ala | Ala | Ser | Leu | Glu | Ala | Glu | His | Gln | Ala | Ile | Val |
|     |     |     | 20  |     |     |     | 25  |     |     |     |     | 30  |     |     |     |
| Arg | Asp | Val | Leu | Ala | Ala | Gly | Asp | Phe | Trp | Gly | Gly | Ala | Gly | Ser | Val |
|     |     | 35  |     |     |     | 40  |     |     |     |     |     | 45  |     |     |     |
| Ala | Cys | Gln | Glu | Phe | Ile | Thr | Glu | Leu | Gly | Arg | Asn | Phe | Gln | Val | Ile |
|     |     | 50  |     |     |     | 55  |     |     |     |     |     | 60  |     |     |     |

Tyr Glu Gln  
45

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 58 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Ser | Arg | Phe | Met | Thr | Asp | Pro | His | Ala | Met | Arg | Asp | Met | Ala |
| 1   |     | 5   |     |     |     |     | 10  |     |     |     |     |     | 15  |     |     |
| Gly | Arg | Phe | Glu | Val | His | Ala | Gln | Thr | Val | Glu | Asp | Gln | Ala | Arg | Arg |
|     |     | 20  |     |     |     |     | 25  |     |     |     |     |     | 30  |     |     |
| Met | Trp | Ala | Ser | Ala | Gln | Asn | Ile | Ser | Gly | Ala | Gly | Trp | Ser | Gly | Met |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |
| Ala | Gln | Ala | Thr | Ser | Leu | Arg | Thr | Met | Thr |     |     |     |     |     |     |
|     |     | 50  |     |     |     |     | 55  |     |     |     |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 94 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Thr | Ile | Asn | Tyr | Glu | Phe | Gly | Asp | Val | Asp | Ala | His | Gly | Ala | Met |
| 1   |     | 5   |     |     |     |     | 10  |     |     |     |     |     | 15  |     |     |
| Ile | Arg | Ala | Gln | Ala | Ala | Ser | Leu | Glu | Ala | Glu | His | Gln | Ala | Ile | Val |
|     |     | 20  |     |     |     |     | 25  |     |     |     |     |     | 30  |     |     |
| Arg | Asp | Val | Leu | Ala | Ala | Gly | Asp | Phe | Trp | Gly | Gly | Ala | Gly | Ser | Val |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |
| Ala | Cys | Gln | Glu | Phe | Ile | Thr | Gln | Leu | Gly | Arg | Asn | Phe | Gln | Val | Ile |
|     |     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |
| Tyr | Glu | Gln | Ala | Asn | Ala | His | Gly | Gln | Lys | Val | Gln | Ala | Ala | Gly | Asn |
|     |     | 65  |     |     |     | 70  |     |     |     | 75  |     |     |     | 80  |     |
| Asn | Met | Ala | Gln | Thr | Asp | Ser | Ala | Val | Gly | Ser | Ser | Trp | Ala |     |     |
|     |     |     |     |     |     | 85  |     |     |     |     |     | 90  |     |     |     |

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Met | Leu | His | Gly | Val | Arg | Asp | Gly | Leu | Val | Arg | Asp | Ala | Asn | Asn |
| 1   |     |     | 5   |     |     |     |     | 10  |     |     |     |     |     | 15  |     |
| Tyr | Glu | Gln | Gln | Glu | Gln | Ala | Ser | Gln | Gln | Ile | Leu | Ser | Ser |     |     |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Thr | Ile | Asn | Tyr | Gln | Phe | Gly | Asp | Val | Asp | Ala | His | Gly | Ala | Met |
| 1   |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |     |
| Ile | Arg | Ala | Gln | Ala | Gly | Leu | Leu | Glu | Ala | Glu | His | Gln | Ala | Ile | Ile |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |
| Arg | Asp | Val | Leu | Thr | Ala | Ser | Asp | Phe | Trp | Gly | Gly | Ala | Gly | Ser | Ala |
|     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |
| Ala | Cys | Gln | Gly | Phe | Ile | Thr | Gln | Leu | Gly | Arg | Asn | Phe | Gln | Val | Ile |
|     |     |     | 50  |     |     |     | 55  |     |     |     | 60  |     |     |     |     |
| Tyr | Glu | Glu | Ala | Asn | Ala | His | Gly | Gln | Lys | Val | Gln | Ala | Ala | Gly | Asn |
| 65  |     |     |     | 70  |     |     |     |     |     | 75  |     |     |     | 80  |     |
| Asn | Met | Ala | Gln | Thr | Asp | Ser | Ala | Val | Gly | Ser | Ser | Trp | Ala |     |     |
|     |     |     | 85  |     |     |     |     |     |     | 90  |     |     |     |     |     |

## (2) INFORMATION FOR SEQ ID NO:22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 69 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```

Ala Arg Arg Met Trp Ala Ser Ala Gln Asn Ile Ser Gly Ala Gly Trp
1           5           10           15
Ser Gly Met Ala Gln Ala Thr Ser Leu Asp Thr Met Ala Gln Met Asn
20           25           30
Gln Ala Phe Arg Asn Ile Val Asn Met Leu His Gly Val Arg Asp Gly
35           40           45
Leu Val Arg Asp Ala Asn Asn Tyr Glu Gln Gln Glu Gln Ala Ser Glu
50           55           60
Gln Ile Leu Ser Ser
55

```

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```

Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met
1           5           10           15
Ile Arg Ala Gln Ala Gly Leu Leu Gln Ala Gln His Gln Ala Ile Ile
20           25           30
Arg Asp Val Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala Gly Ser Ala
35           40           45
Ala Cys Gln Gly Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile
50           55           60
Tyr Gln Gln Ala Asn Thr His Gly Gln Lys Val Gln Ala Ala Gly Asn
65           70           75           80
Asn Met Ala Gln Thr Asp Ser Ala Val Asn Ser Ser Trp Ala
85           90

```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```

Gly Met Ala Glu Ala Thr Ser Xaa Asp Thr Met Thr Gln Met Asn Gln
1             5             10             15
Ala Phe Arg Asn Ile Val Asn Met Leu His Gly Val Arg Asp Gly Leu
20             25             30
Val Arg Asp Ala Asn Xaa Tyr Glu Gln Gln Glu Gln Ala Ser Gln Gln
35             40             45
Ile Leu Ser Ser
50

```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```

Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met
1             5             10             15
Ile Arg Ala Gln Ala Gly Ser Leu Glu Ala Glu His Gln Ala Ile Ile
20             25             30
Ser Asp Val Leu Thr Ala Ser Asp Phe Tyr Gly Gly Ala Gly Ser Ala
35             40             45
Ala Cys Gln Gly Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Xaa
50             55             60
Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn
65             70             75             80
Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Tyr Ala
85             90

```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 98 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```

Met Thr Ser Arg Phe Met Thr Asp Pro His Ala Met Arg Asp Met Ala
 1             5             10             15
Gly Arg Phe Glu Val His Ala Gln Thr Val Glu Asp Glu Ala Arg Arg
      20             25             30
Met Trp Ala Ser Ala Gln Asn Ile Ser Gly Ala Gly Trp Ser Gly Met
 35             40             45
Ala Glu Ala Thr Ser Leu Asp Thr Met Ala Gln Met Asn Gln Ala Phe
 50             55             60
Arg Asn Ile Val Asn Met Leu His Gly Val Arg Asp Gly Leu Val Arg
 65             70             75             80
Asp Ala Asn Asn Tyr Glu Gln Gln Gln Gln Ala Ser Gln Gln Ile Leu
      85             90             95
Ser Ser

```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```

Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met
 1             5             10             15
Ile Arg Ala Asn Ala Gly Leu Leu Gln Ala Glu His Gln Ala Ile Ile
 20             25             30
Ser Asp Val Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala Gly Ser Ala
 35             40             45
Ala Cys Gln Gly Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile
 50             55             60
Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn
 65             70             75             80
Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala
      85             90

```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```

Arg Phe Glu Val His Ala Gln Thr Val Glu Asp Glu Ala Arg Arg Met
 1           5           10           15
Trp Ala Ser Ala Gln Asn Ile Ser Gly Ala Gly Trp Ser Gly Met Ala
 20           25           30
Asn Ala Thr Ser Leu Asp Thr Met Ala Gln Met Asn Gln Ala Phe Arg
 35           40           45
Asn Ile Val Asn Met Leu His Gly Val Arg Asp Gly Leu Val Arg Asp
 50           55           60
Ala Asn Asn Tyr Glu Gln Gln Gln Ala Ser Gln Gln Ile Leu Ser
 65           70           75           80
Ser

```

## (3) INFORMATION FOR SEQ ID NO:29:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```

Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met
 1           5           10           15
Ile Arg Ala Leu Ala Gly Leu Leu Glu Ala Glu His Gln Ala Ile Ile
 20           25           30
Ser Asp Val Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala Gly Ser Ala
 35           40           45
Ala Cys Gln Gly Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile
 50           55           60
Tyr Glu Gln Ala Asn Ala His Gly Glu Lys Val Gln Ala Ala Gly Asn
 65           70           75           80
Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala
 85           90

```

## (2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

|     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gln | Gln | Gln | Ala | Ser | Gln | Gln | Ile | Leu | Ser | Ser |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 94 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Thr | Ile | Asn | Tyr | Gln | Phe | Gly | Asp | Val | Asp | Ala | His | Gly | Ala | Met |
| 1   |     |     | 5   |     |     |     |     | 10  |     |     | 15  |     |     |     |     |
| Ile | Arg | Ala | Gln | Ala | Gly | Leu | Leu | Glu | Ala | Glu | His | Gln | Ala | Phe | Ile |
|     | 20  |     |     |     |     |     |     | 25  |     |     | 30  |     |     |     |     |
| Arg | Asp | Val | Leu | Thr | Ala | Ser | Asp | Phe | Trp | Gly | Gly | Ala | Gly | Ser | Ala |
|     | 35  |     |     |     |     |     | 40  |     |     |     | 45  |     |     |     |     |
| Ala | Cys | Gln | Gly | Phe | Ile | Thr | Gln | Leu | Gly | Arg | Asn | Phe | Gln | Val | Ile |
|     | 50  |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |     |
| Tyr | Glu | Gln | Ala | Asn | Ala | His | Gly | Gln | Lys | Val | Gln | Ala | Ala | Gly | Asn |
| 65  |     |     |     | 70  |     |     |     |     |     | 75  |     |     |     | 80  |     |
| Asn | Met | Ala | Gln | Thr | Asp | Ser | Ala | Val | Gly | Ser | Ser | Trp | Ala |     |     |
|     |     |     | 85  |     |     |     |     |     |     | 90  |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 99 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ser | Phe | Val | Thr | Thr | Gln | Pro | Glu | Ala | Leu | Ala | Ala | Ala | Ala | Ala |
| 1   |     |     | 5   |     |     |     | 10  |     |     | 15  |     |     |     |     |     |
| Asn | Leu | Glu | Gly | Ile | Gly | Thr | Thr | Met | Asn | Ala | Gln | Asn | Ala | Ala | Ala |